Bolt Beranek and Newman Inc.

LEVEL

Report No. 4724

AD A109491

Display-Control Compatibility in 3-D Displays 1: Effects of Orientation

A.W.F. Huggins and David J. Getty

November 1981

Prepared for:
Engineering Psychology Programs
Office of Naval Research

82 01 009

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER BBN Report No. 4724	2. GOVT ACCESSION NO. D-A109491	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Display-Control Compatibility in 3-D Displays 1: Effects of Orientation		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER BBN Report No. 4724
7. AUTHOR(s) A. W. F. Huggins and David J. Getty		8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0750
9. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Mass 02238		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 196-166
11. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Programs, ONR 800 North Quincy Street, Code 442 Arlington, VA 22217		12. REPORT DATE November 15, 1981 13. NUMBER OF PAGES 74
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

For Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Display-Control Compatibility
3-D Display
SpaceGraph
Reaction Time
Decision Strategy
Orientational Mismatch
Mental Rotation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The recent development of true volumetric displays, such as SpaceGraph, has raised questions about display-control relationships that cannot be answered from earlier work with 2-D displays. We measured the speed and accuracy of simple control decisions when the displayed object was presented in orientations rotated away from congruence with the control orientation. Reaction times were measured for identifying

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

the marked face of a static cube, presented with SpaceGraph, as a function of (a) how much the cube image was rotated away from congruence with fixed physical cube on which the observer responded, and (b) which dimension of the display corresponded to the rotation axis. The following results were obtained: (1) The fastest and most accurate identifications occurred when there was a simple direct spatial mapping between the display and the control for all orientations presented. Decision times made with this strategy were relatively unaffected by orientation. (2) Marked display-control incompatibility was found under some specific conditions; and (3) The shape of the decision time functions suggest that decision strategies involved both propositional coding and mental rotation.

Table of Contents

Table of Figures

INTRODUCTION

Recent developments in 3-D display technology for computers have resulted in devices with immediate and obvious applications in areas such as remote control and manipulation, and in monitoring many-dimensional events such as air traffic patterns. They also point up gaps in our present understanding of display-control compatibility issues, which up till now have usually been concerned with 2-dimensional displays. The work described in this report represents the first phase of a project to explore display-control compatibility issues with true 3-D displays, identify problem areas, and suggest and test possible remedies.

Earlier research has distinguished between several ways for displays and controls to be incompatible. Fitts and Seeger (1953) addressed the problem under the title of Stimulus-Response Compatibility. They designed three different displays for indicating one of the directions corresponding to the eight major compass points, and three sets of controls by which the observer specified the displayed direction. Each set of controls was a close spatial analog of one of the displays. Choice reaction times were measured for each of the nine possible combinations of display and control. Response times showed a major interaction between the display mode and the response mode, reflecting the fact that responding was much faster and more accurate when the response set was a direct spatial mapping of the stimulus set

than otherwise. This and later studies led to the conclusion that how a set of responses is assigned to a set of stimuli is a major determinant of response speed and accuracy, in addition to the effects due to the particular stimulus and response set. To avoid circularity, compatibility was often defined in terms of responses that are the most "natural" or obvious, or expected by naive observers; that is, the responses that correspond to the population stereotype. Where spatial displays are concerned, of course, a direct spatial correspondence is superior to any other mapping.

Unnaturalness, or incompatibility, can arise from several sources: first, any departure from the mapping suggested by the natural geometry of the stimuli and responses will decrease speed and accuracy. Second, if the rule that maps stimuli into responses is complicated, decoding the rule will take time, and the response will be less than optimally rapid. For example, if the stimulus is an integer displayed on a CRT, an even/odd response can be made much more rapidly (i.e. is more compatible) to the number itself than to the sum of its digits, or to the largest square smaller than the number. A further type of incompatibility arises from response competition: if cards must be sorted into six piles according to the number of symbols that appear on their faces, sorting is much slower if the symbols are themselves the digits 1-6 than if they are meaningless shapes (Morton, 1969). This task is a spatial analog of the well-known Stroop test (Stroop, 1935), which involves naming the color of

the ink in which each word in a list is printed. Performance is considerably degraded if the words are themselves the names of _different_ colors. Response competition typically occurs when there is, in the set of possible responses, a natural or stereotyped response to a particular stimulus, but the mapping rules decree that a _different_ response be made.

Thus three conditions can be distinguished. In the first, a stimulus is presented but _none_ of the available responses seems particularly appropriate, and considerable processing of the stimulus must occur before the response can be selected (incompatibility due to mapping complexity). In the second condition, exactly _one_ of the available responses seems particularly appropriate, and this is also the correct response (compatible response). In the third condition, again one of the available responses seems appropriate, but this is _not_ the correct response, and must be inhibited until the processing that specifies the correct response is complete (incompatibility due to response competition). The earlier an inappropriate response becomes available, the longer it must be suppressed, and the more interference it causes. In support of this interpretation, Sanders (1967) has shown that no interference with the incompatible correct response is caused by requiring the observer to make the faster compatible INcorrect response as well as the slower correct response. Other supporting evidence can be found in Woods (1974). Woods showed that irrelevant variation in a late-available phonetic discrimination (spoken _b_ versus _g_) had no

effect on the speed of an early-available auditory discrimination (whether the syllable was spoken with high or low voice pitch). Irrelevant information that became available after the response had been selected did not delay the response; Woods interpreted this as evidence for separate parallel processing at the auditory and the phonetic levels. On the other hand, irrelevant variation in pitch did interfere with the phonetic decision, so irrelevant information that arrived before response selection did interfere with the response. Furthermore, when the variation in the two variables was correlated (e.g. the b was always spoken with low pitch), the distribution of reaction times was well modelled by combining the distributions for the auditory and phonetic decisions to yield a joint distribution for "whichever of the decisions happened to become available first."

In considering display-control compatibility in 3-D displays, we will limit our attention to spatial arrays of stimuli, and responses that represent 1:1 spatial mappings of them. Two types of incompatibility mentioned above can occur in spatial displays. For example, if the six spatial directions (up, down, left, right, towards, and away) are associated with a linear row of six response keys, no assignment is likely to be highly compatible (although those that are random rather than regular may be even less compatible). One possible compatible response set consists of six response keys mounted on the faces of a cube, one key corresponding to the "up" direction and so on. If the directions were mapped onto responses in the opposite

direction (e.g. requiring the bottom key to be pressed when the "up" direction is presented) the responses would be slowed due to response competition.

There appear to be few studies that have looked systematically at incompatibility resulting from 3-D mismatches of orientation. In a study of direction-of-motion stereotypes -- what the majority of a population regard as "natural" -- Mitchell and Vince (1951) found large differences in efficiency between different movements of a control lever for producing a specified movement of a display pointer. For an upward movement of the pointer, the control movements were ordered from most to least effective as follows: upward, forward, right or left, backward, and downward. Error rates ranged from about 5% for the upward movement to about 13% for the downward movement. Warrick (1947) studied population stereotypes in a task that involved turning a rotary control so as to change which in a row of bulbs was lit from one of the end bulbs to the one in the center. When the control knob was mounted on the front face of the box, below the row of lights, observers agreed that clockwise rotation of the knob should move the lit light towards the right. But when the knob was on the side of the box, there was little agreement. Population stereotypes such as these are strongly influenced by how the observers conceptualize the linkage between the control and the pointer.

Other studies have typically considered 2-dimensional displays. For example, Peterson (1965) used a set of four stimulus lights arranged on the lateral and depth axes of a board in front of the observer, each with a concentric response key around it, and studied response times for each of the 24 possible mappings of the four stimuli onto the four responses. Included among the 24 mappings were four that together represented a complete revolution, in four equal steps, around a vertical axis perpendicular to the stimulus and response arrays. That is, in the four mappings, each stimulus light required a response that was displaced clockwise from it by 0, 90, 180, or 270 degrees, respectively. Reaction times averaged 294 ms in the compatible condition, and 421, 360, and 414 ms, respectively, with mismatches of one quarter, one half, and three-quarters of a revolution. Thus the congruent mapping was the most compatible, yielding the shortest response times, and the quarter-revolution mappings were the least compatible. These data, while interesting, are clearly insufficient to predict the effects of arbitrary mismatches of 3-D orientation, such as might occur in a 3-D control task.

Other lines of research that are relevant to the problem at hand concern (a) mental rotation of images and (b) cognitive maps. Cooper and Shepard (1973) presented observers with rotated letters and digits, and asked them to decide whether the letters were normal or backwards (left-right reflected). In the absence of preview of either the identity of the letter or the

orientation in which it would appear, they obtained a reaction time function shaped like an inverted V, with its minimum corresponding to the upright presentation and its maximum to the inverted orientation. Separate preview of the identity and orientation of the stimulus letter allowed observers to prepare a mentally-rotated image of the normal letter, and base their response on a same-different judgment. Cooper and Shepard deliberately chose a letter matching task that would require their observers to perform mental rotations. It is not clear whether operators in a 3-D control task have to compensate for any orientational mismatch by performing mental rotations on the information presented in the 3-D display, before they can select a response, but clearly it is one possible strategy. It is also possible that Peterson's (1965) observers used mental rotation in selecting their responses, although a coding in terms of a set of translation rules seems more likely since this was the only option available for many of the other 20 mappings tested. In any case, the inverted V-shaped function found by Cooper and Shepard is clearly quite different from the M-shaped function found by Peterson. Another difference between the experiments that might contribute to the different results is the orientation of the plane in which the orientational mismatch between display and controls, or between display and mental image, occurred. In Peterson's study, stimulus lights and response keys were mounted on a horizontal board in front of the observer, so "rotation" was about a vertical axis. In Cooper and Shepard's experiment, on

the other hand, since the letters were presented with a tachistoscope, presumably the rotation took place in the vertical plane. This may or may not be an important difference: something akin to mental rotation of an image occurs every time the focus of visual attention is changed, and one could argue that, since most human activities are performed with the head erect, people have much more practice in dealing with mental rotations about a vertical axis than about any other axis, and this might result in differences in the processing strategies used for different axes of rotation.

The study involving cognitive maps (Hintzman, O'Dell, & Arndt, 1981) addresses the foregoing idea rather directly. In effect, their observers were asked to imagine themselves facing in one of the eight major compass directions, and to then "point" in one of the eight directions. Pointing was performed by moving a stylus towards one of eight metal arcs arranged in a circle around a central "home" button. Reaction times suggested that orientational shifts were achieved through mental rotation when the stimulus information was presented in the form of an arrow in a visual display, but not when observers memorized a specific <u>object</u> at each direction, and then answered questions like "if you face the window, where is the chair?" With the visual display, reaction time as a function of <u>response</u> direction relative to the observer was shortest for FRONT and BACK responses, and elevated by a roughly constant amount for all other directions (i.e. the function was M-shaped). However, when

reaction time was plotted as a function of the direction the observer was asked to imagine him/herself facing, the function had an inverted V shape, and preview of this orientation led to a lowering of the reaction times, but no change in the shape of the function. These results are remarkably similar to those obtained by Cooper and Shepard, even though the tasks were quite different, and suggest that the observers rotated an image of the display to a preferred orientation with the arrow pointing straight up, before selecting the pointing response.

The foregoing discussion has raised several questions about how 3-D mismatches of orientation might affect responses in a 3-D control task, and the purpose of the experiments described below was to measure these effects directly. Four separate experiments were performed. In each, a stimulus cube was presented to the observer with the SpaceGraph display, which generated a true volumetric virtual image by having the observer view a time-varying CRT display in a vibrating mirror. One face of the cube was marked, and the observer's task was to decide which face was marked, and press the response key on the corresponding face of a fixed response cube in front of him/her. In the first experiment, the stimulus cube was presented in 24 different orientations arranged at 15 degree (0.26 rad) intervals around the vertical (Y) axis. The second and third experiments were identical to the first, except that the stimulus cube was rotated about its depth (Z) axis, or its lateral (X) axis, respectively. In the fourth experiment, simple rotations were made about any one of the three axes, chosen randomly on each trial.

EXPERIMENT 1

OBSERVERS

The same three observers served in all four experiments. Two were the authors, and the third was the 16-year old male son of one, who was naive with respect to the purposes of the experiments. Two wore glasses, but all had normal stereopsis. A fourth observer, a research assistant at BBN, also served in the first experiment, and gave similar results to the other three. We have not included his data because he was unable to complete more than the first task, for reasons unrelated to the experiment.

APPARATUS

The experimental apparatus is illustrated in Figure 1. The observer sat in a chair with a cubical response box mounted on a shelf in front of him. The stimulus image was presented on the SpaceGraph display, described in detail below. Not shown in the figure are two PDP-11 computers, linked through a 1200 baud serial line. One computer served as the host for the SpaceGraph display, and read in display files and turned the display on and off under the control of the second computer. The second computer controlled the sequence of trials, instructed the SpaceGraph host what files to display and when, timed the responses by monitoring the response keys, and recorded the results of each trial.

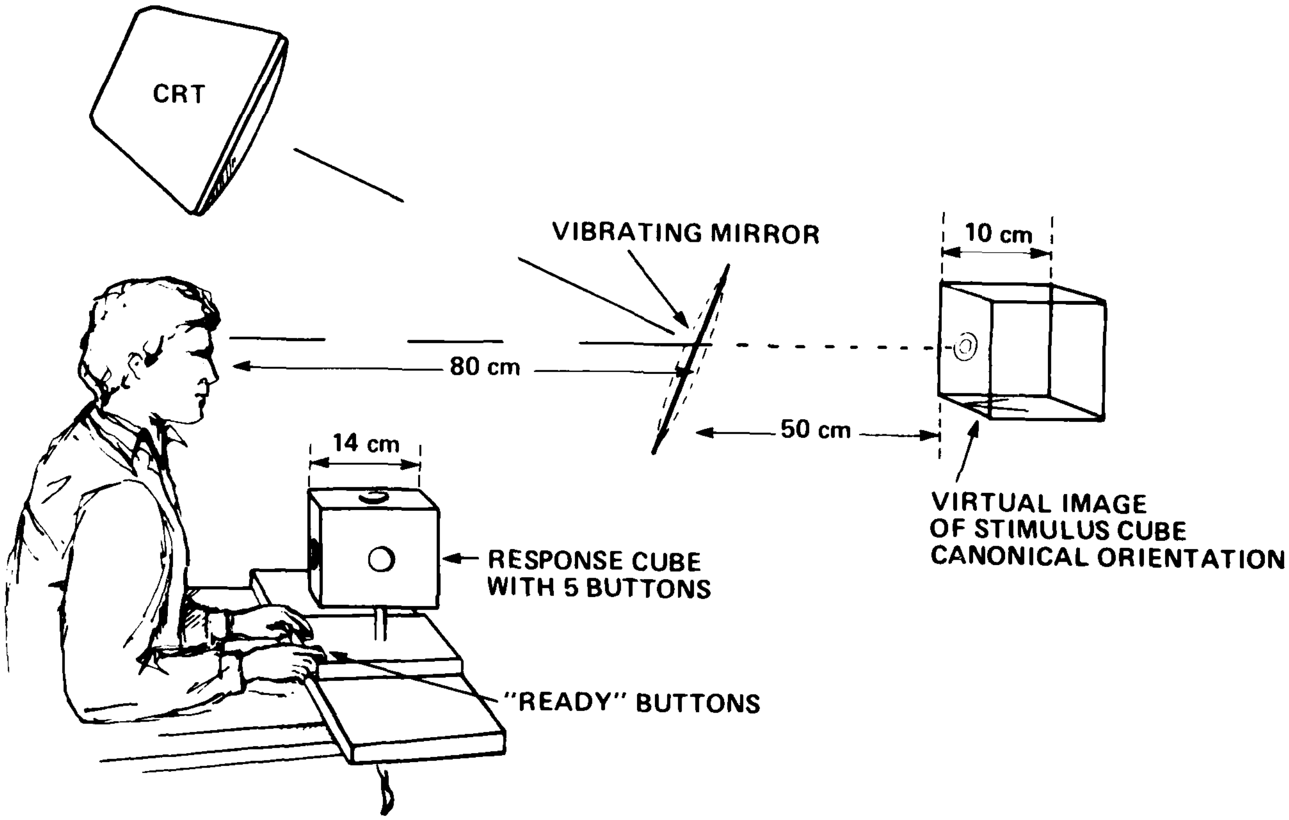

Figure 1: Sketch of observer with response box and SpaceGraph display. The observer viewed the CRT in a vibrating mirror, which generated a virtual image of the stimulus cube behind the mirror.

The Stimulus Image

The displayed image consisted of an outline of a cube (12 edges), with sides about 10 cm long. The head-on orientation of the cube is defined as that in which the vertical axis of the cube corresponds to the gravitational vertical, and the lateral horizontal axis of the cube is perpendicular to the observer's line of sight through the center of the cube. The viewing axis of the display declines at about 10 degrees (0.18 rad) below the horizontal, so the observer looks down slightly on the cube. The cube appeared in any one of 24 different orientations equally spaced at 15 degree (0.26 rad) intervals around the cube's vertical Y-axis, beginning with a 10 degree (0.18 rad) rotation from the head-on position to minimize the problems of inter-point contention inherent in the display, explained in more detail below. To mark the orientation of the cube, a capital letter V was drawn on the bottom face, extending about half way across the face, with its apex nearly touching the front edge. Two small concentric circles, representing a stimulus key or button, appeared at the center of one face (any except the bottom), to mark the face that was to be identified by the observer.

Photographs of the 24 stimulus orientations are shown in Figure 2. The photographs differ from what the observers actually saw in two ways. First, the stimulus images never contained more than one stimulus key at a time, whereas in the figure all of the five stimulus keys appear. Second, it is quite

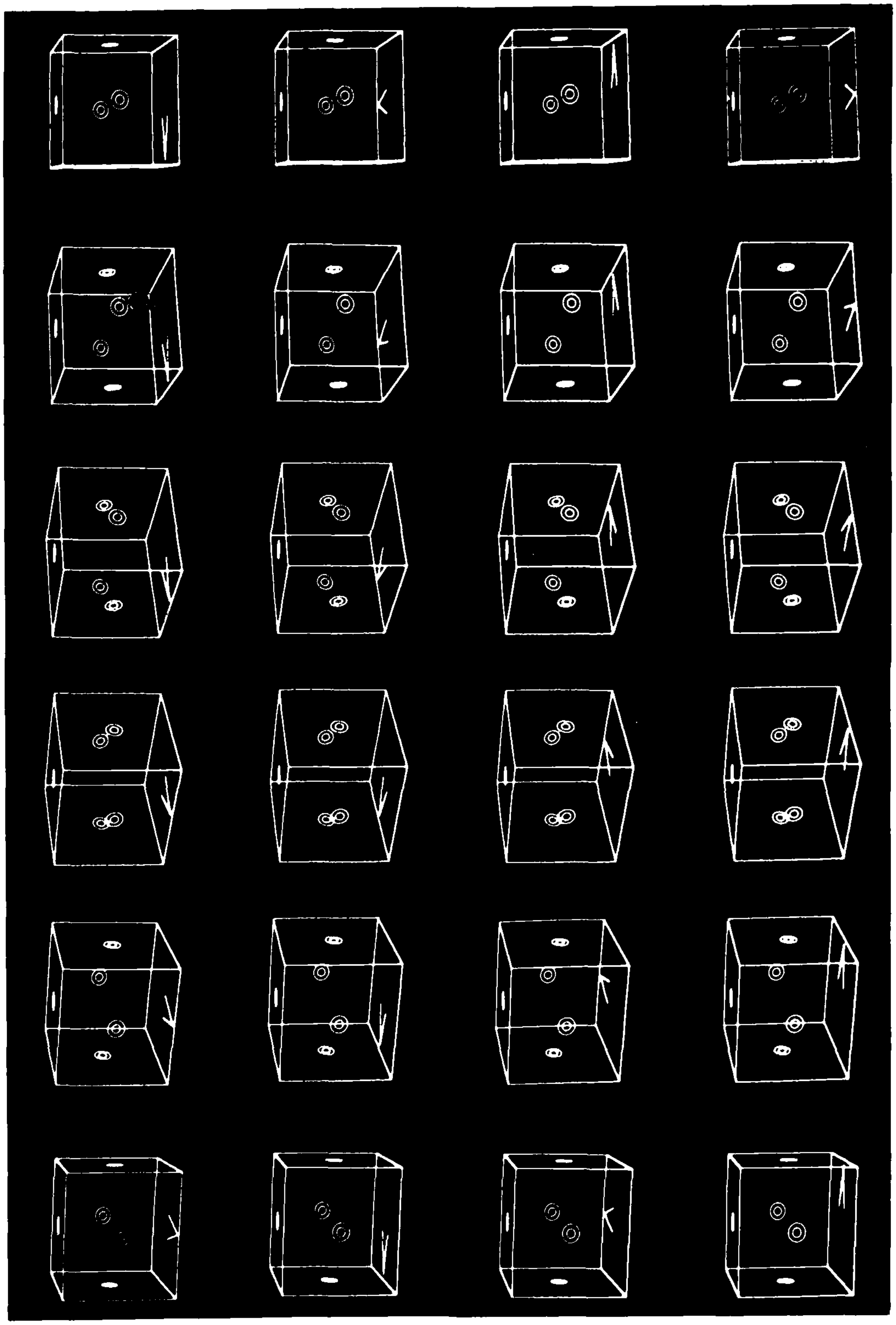

Figure 2: The 24 orientations of the stimulus cube used in Experiment 1, in which rotation was around the cube's vertical or Y-axis. All five stimulus "keys" are shown in each view, to save space. On each trial, the observer saw only one of the five stimulus keys.

impossible to capture in a two-dimensional static photograph the immediacy of the three-dimensional virtual image. Some of the 2-D photos are quite hard to interpret, but this was virtually never true of their 3-D counterparts. The salience of the 3-D percept generated by SpaceGraph is illustrated by how seldom observers experienced depth ambiguity, despite the strong similarity between our stimuli and the Necker cube, a two dimensional image designed to maximize depth ambiguity. The observers reported experiencing depth ambiguity in only about twenty trials out of a total of 49,820 trials on which data were collected.

The Response Box

The response box was a slightly flattened cube of sheet aluminum, 14 cm square on the top and 11.5 cm high, rigidly mounted by a vertical stalk on a shelf fixed to the observer's chair (see Figure 1). Mounted in the center of each side of the cube was a round flush-mounted plexiglass response key 6 cm in diameter. Two "ready" keys, similar to the other response keys, were mounted about 2 cm above the shelf, and slightly outside and nearer to the observer than the front corners of the response cube. Each response key operated a snap switch if any part of its surface was moved inwards by about 1 mm (or downwards for the ready keys). The force to operate a switch was about half a pound (2.2 N).

The SpaceGraph Display

Physical Description: The display used for presenting the stimuli was the prototype of a new genre of true space-filling display systems developed at Bolt Beranek and Newman Inc., called SpaceGraph (TM). SpaceGraph consists of a computer-driven CRT display that is viewed in a mirror, which makes the CRT appear to be behind the mirror. The mirror is a circular coated sheet of flexible plastic, mounted by a circular hinge on a high-quality low-frequency loudspeaker. When the loudspeaker is excited by a 30 Hz sine wave, the mirror flexes spherically, and behaves optically like a lens of continuously varying focal length. As a result, the virtual image of the CRT face moves cyclically towards and away from the viewer at 30 Hz. A peak mirror vibration amplitude of about 0.3 cm yields a cyclic displacement of the image of the CRT face of about 20 cm. A point can be displayed anywhere in the resulting display volume that is swept out by the virtual image of the CRT face: the X and Y coordinates of the image are addressed as the X and Y coordinates of the CRT (with a scaling correction to compensate for the changes in magnification of the image), and the Z, or depth, dimension is addressed by the instant in the display cycle at which the point is flashed on the CRT.

Properties of SpaceGraph: A SpaceGraph image consists of luminous points (and lines and surfaces made up of arrays of points) that seem to float in a 20 cm cube behind the mirror and

about 1.5 meters from the observer. Since the image is refreshed 30 times per second, it is free of flicker (although naive viewers sometimes see the image "jump" if they make a sudden head movement). The image can be viewed from positions up to about 30 cm off axis of the 40 cm diameter mirror. In its point and line drawing mode, SpaceGraph is capable of displaying between 5,000 and 10,000 points within the volume, each at any of 8 different brightness levels, which is sufficient for fairly complex images.

SpaceGraph has several properties that make it unique as a display. First, the virtual image fills a real three-dimensional volume. Because the points in the virtual image are truly at different depths from the observer, head movements cause the relative positions of the points, and the perspective of objects, to change dynamically exactly as would those of real objects. The resulting dynamic parallax, which is crucial for true depth perception, is very cumbersome to achieve in other 3-D displays such as stereo TV, since it requires mechanically slaving the pair of TV cameras to movements of the observer's head and each of his eyes (to replicate convergence movements as well as changes of gaze). Then there are further constraints to be met: for instance, the horizontal plane of the two TV images and the horizontal plane of the viewer's eyes must coincide, or the viewer may be unable to fuse the images. SpaceGraph, on the other hand, requires no special equipment, and head movements enhance the depth percept rather than the reverse.

A second property of SpaceGraph that makes it different is that, since objects are represented in the display as arrays of luminous points floating in space, objects are transparent. That is, instead of a farther object being partly occluded by a nearer intervening object, the farther object can be seen through the nearer. It is not obvious whether transparency is an advantage or a disadvantage. When an object is being manipulated towards a target, it probably helps if the object is transparent, so that the target can be seen through it. On the other hand, the fundamental visual distinction between figure and ground may be blurred by the transparency, because the ground is always visible <u>through</u> the figure, and may therefore act to obscure it. Thus increasing the background detail, to make it easier for the operator to orient himself in the display, may have the undesired effect of making it harder to pick out the objects of primary interest from the background. The usual way of handling this problem in 2-D displays, hidden-line removal, is not appropriate for true 3-D displays, since the point where a particular line is occluded depends on the observer's position. Therefore, hidden line removal can be achieved only by fixing the observer's head, thus removing movement parallax cues to depth.

Inherent in the principle of operation of SpaceGraph is a constraint that limits its ability to draw lines truly normal to the viewing axis. Because the mirror in which the display is viewed is vibrating, and because it takes a finite time to display each point, no two points can be displayed at <u>exactly</u> the

same depth (except at the extreme front and back of the display volume, or unless points are plotted on the foreswing of the mirror as well as the backswing, an option not currently implemented). When two points contend for the same depth plane, one can be discarded, or both can be plotted but at (imperceptibly) different depths (adjacent depth planes are less than 0.006 cm apart). Lines are composed of sequences of points, so a line plotted perpendicular to the viewing axis presents severe contention problems. There are two possible solutions. One is to tilt the line slightly away from the normal. There is, in fact, a direct relationship between the length and required density of a line and the minimum tilt at which this density can be achieved. In order to plot long lines with very small tilt (e.g. less than one degree or 0.018 rad) the lines must be very sparse (i.e. few points per cm). With this solution to the contention problem, one end of the line will appear at a slightly greater depth than the other. This solution is not appropriate when the line is part of the front face of a cube, however, since both ends of the line are constrained by having to intersect the other lines that define the cube's corners. The second solution is to plot the ends of the line as nearly as possible in the same depth plane, and fill in the middle of the line later. As a result, the middle of the line sags slightly away from the viewer in depth. Thus, in this solution, contention is resolved by relaxing the requirement that the line be exactly straight. In practice, both of these solutions lead to distortions of the

image that are imperceptible. Nevertheless, we avoided images that required lines to be drawn even approximately perpendicular to the viewing axis.

When a line is plotted with sufficient tilt that contention does not occur between successive points in the line, contention arises only with other lines crossing the same depth planes, or with points lying in these planes. The dependence of contention on the orientation of the line can lead to substantial differences in image quality and perceived brightness as a function of the orientation of the lines in the display space. A further problem arises when, as in the present study, a marker must be added to one of the faces of the cube shown in the display. The points in the marker may contend unpredictably with the points defining the cube edges, with the result that the edges may have gaps in them. Gaps add an unwanted complexity to the observer's task, since sometimes they will assist the observer by providing extra information about which cube face is marked, and at other times they will distract him, since the positions of the gaps will vary as a function of the overall orientation of the cube. To avoid changes in the cube-image that were correlated with which of its faces was marked, the display file for a particular cube orientation contained the points necessary for showing all of the five stimulus keys (it is these files that appear in Figure 2). Before the image was presented to the observer in an experimental trial, all but one of the stimulus keys were removed from the image by using an instruction

code that failed to unblank the CRT for those points. This ensured that imperfections in the cube image could not help the observer identify the marked face. The imperfections in the image due to contention were in turn minimized (a) by avoiding any orientation of the cube image that required lines to be drawn perpendicular to the display's depth axis; (b) by reducing the density with which all lines were plotted, and (c) by reducing the mirror amplitude until the display volume was just deep enough to contain the 10 cm stimulus cube in any orientation (thus reducing the distance between adjacent depth planes). Since the viewing axis of SpaceGraph is inclined at 10 degrees (0.18 rad) to the horizontal, rotating the cube image forwards by this amount both eliminated the contention problem for "vertical" lines in the image, and also made the cube's vertical axis coincide with the gravitational vertical, which was desirable for other reasons. To eliminate contention in horizontal lines, we define the canonical (or zero-mismatch) orientation of the cube to be one in which the cube image was rotated 10 degrees (0.18 rad) clockwise looking down the vertical axis, away from the head-on orientation. The 24 orientations were spaced at 15 degree (0.26 rad) intervals from this canonical orientation. Reducing the line density and avoiding true head-on orientations, between them, greatly reduced both the undesired variation of brightness and the salience of gaps in the edges to a point where they were imperceptible. We mention these problems with SpaceGraph to document the care we took to ensure that our

stimuli were what we required. It took several preliminary attempts to find a display specification that met our requirements. It is important to stress that these problems will not be of concern in most practical applications of SpaceGraph.

Calibration of the display: Since the image viewed by the observer is a virtual image, calibrating the display raised some difficulties. The purpose of calibrating the display was less to establish the exact dimensions of the stimulus cube and the absolute brightness of the lines than to ensure that the settings could be recaptured from one session to the next, so that identical images would be seen on different days. There is no direct way of measuring the true depth of the virtual image. Placing a ruler in the virtual image (behind the mirror) is possible, but measuring the image requires viewing the cube with the left eye and the ruler with the right, and making the measurement at arm's length, so it is probably not very accurate. Furthermore, the mirror may not be truly spherical near its edges, where the image must be viewed if the right eye is to see past the edge of the mirror to the ruler, and the resulting distortion of the cube image may introduce further error.

We designed a test pattern for adjusting the mirror amplitude before each session. It consisted of five lines running from the four corners and the center of a square at the back of the display volume to the opposite corners and the center

of a similar square at the front of the display volume. When the mirror amplitude was set too low, the display had insufficient depth, and the four slanted lines exited the display behind the mirror and it was not possible to sight along them. If the amplitude was set too high, the lines appeared to intersect the mirror inside its outside edge, and it was possible to sight along them. When the amplitude was correct, each of the four lines intersected the edge of the mirror. The calibration display that had the right properties was selected by trial and error: first the mirror amplitude was varied to find where a cube image became noticeably too shallow or deep. Then the amplitude was set midway between these values, and a series of calibration displays was shown, and the one whose lines intersected the edge of the mirror was selected. The same calibration display served to ensure that the image was displayed symmetrically with respect to the phase of the mirror. This is important, since otherwise lines that are supposed to be straight may be slightly curved.

PROCEDURE

Each experimental session consisted of an initial block of 25 warm-up trials, followed by four blocks of 120 experimental trials each, and a final "calibration" block of 25 trials. In the warmup block, each orientation of the stimulus cube appeared once, except that the canonical orientation appeared twice. Each button was presented five times, and both orientations and buttons occurred in an irregular order. Each block of

experimental trials contained one presentation of each of the 120 different combinations of the five buttons with the 24 orientations. The 120 stimulus images were presented in a different random order (without replacement) in each block of each session, but all observers saw the same sequence of trials. The final calibration block used only the canonical orientation, and each button was presented five times in an irregular order.

The observer indicated that he was ready for a trial by placing his left and right hands on the two "ready" keys. Two seconds later, the cube appeared in front of him, in one of the 24 orientations and with one of the five stimulus buttons showing. The observer held down the ready keys until he had decided which face of the stimulus cube was marked with the stimulus button, and then released the ready keys and quickly pressed the key at the center of the appropriate face on the response cube in front of him. All three of our observers were right-handed, and chose to press the LEFT response key with their left hands, and all other response keys with their right hands. The stimulus cube and button remained visible until the correct response key was pressed, so feedback of an error was provided by failure of the image to disappear. Pressing the ready keys again began the next trial, so strictly the task was self paced. In practice, however, observers did not pause between making a response and pressing the ready buttons for the next trial, and the observer knew when to expect the next image to appear. Observers were allowed to pause between trials, but were asked

not to pause more than briefly except between blocks. The observers were told to respond as fast as they could without making any errors. At the end of each block of trials, a message was shown in the display informing the observer how many errors he had made in the preceding block. (This feature was added during the first experiment, in response to the high error rate of one observer: see further below.) Each observer served for ten sessions, with a maximum of two experimental sessions in a day, and a rest of at least 30 minutes intervening. Each session lasted 30-40 minutes.

Two times were recorded for each trial: the "reaction" time, from the presentation of the stimulus cube to the release of the ready key by the hand that then made the response, and the "movement" time, from the release of the ready key to the depression of the response key.

RESULTS

Although ten sessions of data were collected with each observer, only the data from sessions 2 to 6 will be discussed below. The first session was discarded to reduce the effects of familiarization, and the last four sessions were discarded to make the results of the four experiments directly comparable (only six sessions were run on Experiments 2-4). Our interpretations would not be materially affected if we had included all the data.

Errors

The three observers made a total of 134 errors in the 7200 trials in sessions 2-6, for an average rate of 1.9%. A few more responses were lost due to hardware failures. Two of the observers made very few errors: one made 21 in 2400 trials and the other made 32, for overall rates of 0.9% and 1.3% respectively. The third observer (the high-school student) made 79 errors, for an average of 3.3%. Data analysis was slightly delayed behind data collection, with the result that the higher error rate was not detected until the eighth session. The experimental procedure was changed so that a text display was shown to the observer at the end of each block of trials, stating how many errors had occurred. The third observer was re-instructed, and thereafter his error rate fell to the same level as the others, about 1%. (In fact, in the 960 trials of sessions 9 and 10 he made a total of 6 errors, for an average rate of 0.6%. This was accomplished at the cost of an increase in the mean reaction time of 80 ms, averaged over all 5 responses and 24 orientations, compared with a typical variability, from one pair of sessions to the next for a single observer, of about 30 ms.)

Confusions between LEFT and RIGHT responses accounted for 83% of the errors, with 52% being LEFT responses to RIGHT stimuli, and 31% in the reverse sense. Only 8 of the total of 134 errors involved NEAR/FAR confusions, and 14 (all from the

third observer) involved quarter-revolution errors (e.g. NEAR response to LEFT stimulus). Three quarters of the errors were made in orientations more than one quarter revolution away from the head-on position. No errors were made to the TOP stimulus, and the TOP response was never made to any other stimulus. Trials in which errors occurred were not included in further analyses.

Smoothing the data

In each block of data, there were a few responses that were obviously unreasonably long. Sometimes these were due to momentary inattention by the observer, and sometimes (movement times only) the observer's response movement failed to hit the intended response key, and the response had to be repeated. We tried several ways of screening the data to remove these atypical responses, but each method either had an insignificant effect, or removed too much data under some conditions, or both. Furthermore, the implications screening has for statistical analyses is not clear. Therefore, the data we present have not been screened. To reduce the noise in the plotted data, we applied boxcar smoothing to each point before plotting it. In boxcar smoothing, a plotted point represents the average of the true data point for that abscissa value with the two immediately adjacent values.

Movement times

Two intervals were measured for each trial: the reaction time, from the presentation of the stimulus to the release of the ready buttons, and the movement time, from the release of the ready buttons to the pressing of the response key. Mean movement times pooled across observers are plotted for each response in Figure 3 as a function of the orientation of the stimulus cube. In all the plots, the abscissa values represent rotations from the head-on orientation. Thus, the canonical orientation is represented by the data points immediately to the right of the vertical dotted lines at zero rotation. Data points for one quarter revolution in each direction from the head-on view are duplicated at the left and right sides of the plot, for the sake of continuity and to make the symmetry of the minimum at zero rotation more apparent.

Movement times varied slightly for the five different responses, but they varied very little as a function of orientation of the stimulus cube. The movement times for the TOP and FAR responses were about 100 ms longer than the other three responses, presumably reflecting the longer and less direct movements involved. For the TOP, NEAR and FAR responses the functions are almost flat, showing that movement times were essentially independent of the orientation of the stimulus cube. The LEFT and RIGHT responses appear to be slightly elevated for orientations more than one quarter revolution from the head-on

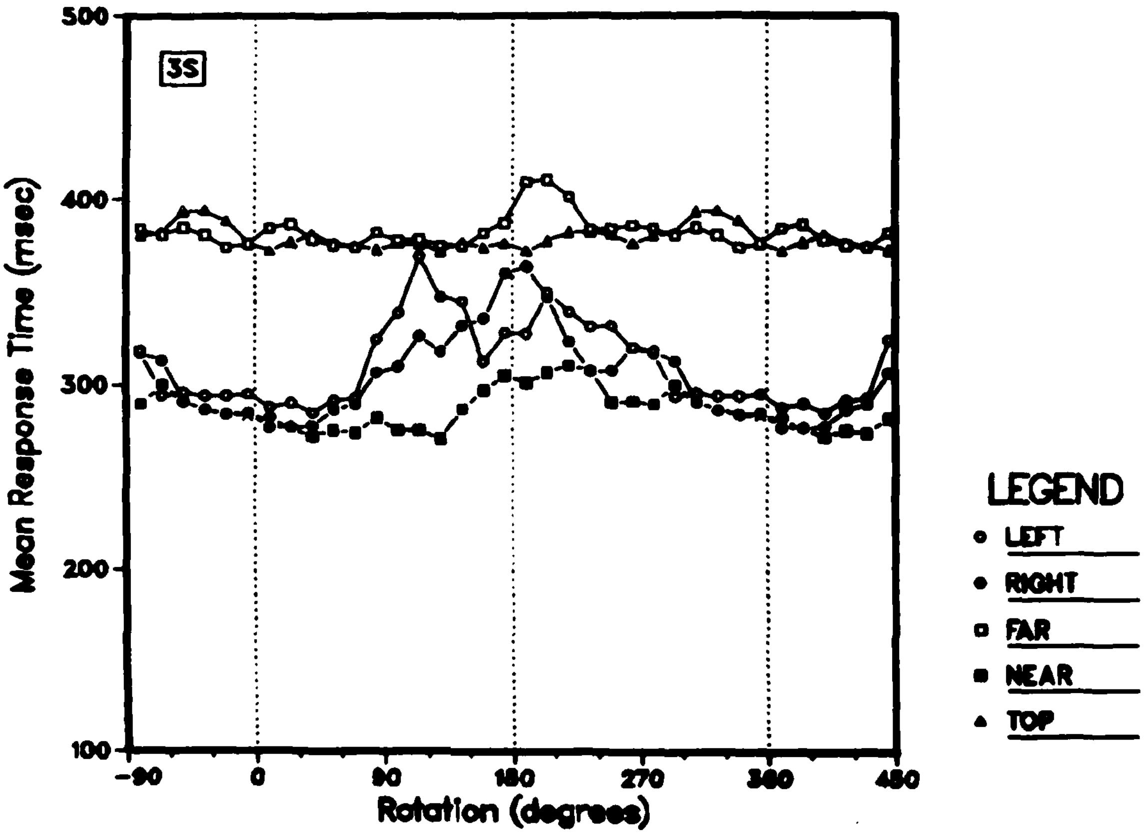

Figure 3: Mean movement times in the Y-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

view. The elevated values were due to the third observer (the high-school student), who made a significant number of errors during sessions 1-8. The elevated values are due to some "leakage" from the peaks of the reaction time functions (see below and Figure 4). After session 8, as mentioned above, this observer was re-instructed. Thereafter his error rate dropped and the leakage disappeared. The leakage occurred because the observer sometimes removed his hands from the ready keys before he had decided on his response, with the result that the movement time included some time spent choosing the correct response.

Reaction times

Mean reaction times pooled across observers are plotted for each of the five responses as a function of stimulus cube orientation in Figure 4. As mentioned above, although ten sessions of data were collected, only the data from sessions 2 to 6 for each observer are plotted in Figure 4. Very similar plots are obtained from sessions 1-10, 2-10, 1-6, and 2-6, and our interpretations would not be materially affected if we had made a different selection.

The functions shown in Figure 4 fall into three groups: that for the TOP responses; those for the NEAR and FAR responses; and those for the LEFT and RIGHT responses. The function for the TOP responses is essentially flat, showing that this reaction took about 740 ms, and was not affected by the orientation of the

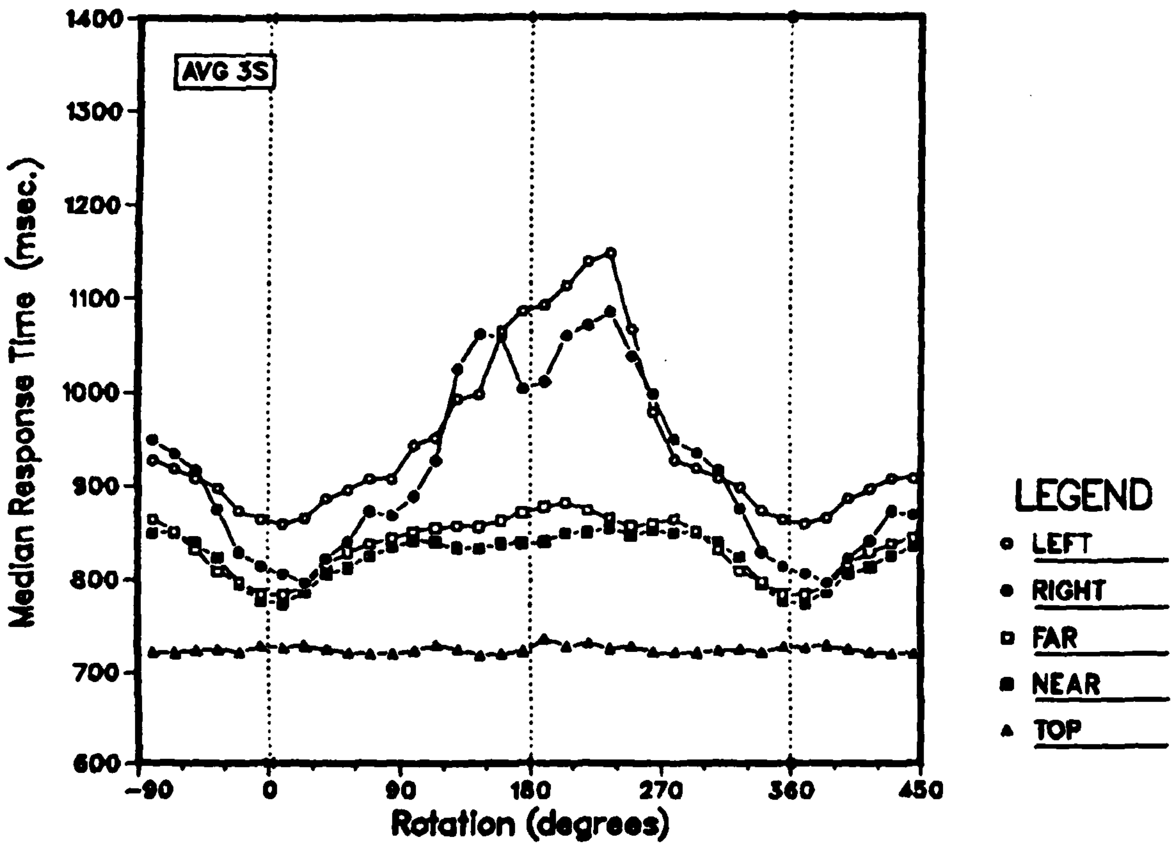

Figure 4: Mean reaction times in the Y-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

stimulus cube. The functions for the NEAR and FAR responses exhibit plateaus. The reaction time is lowest (about 800 ms) at the head-on orientation, and rises roughly linearly with rotation away from the head-on orientation until the plateau is reached for rotations more than about a quarter revolution from the head-on orientation. The plateau is at about 860 ms for the NEAR reaction and 890 ms for the FAR reaction. The third group, the LEFT and RIGHT functions, exhibit much more dramatic effects of orientation. The LEFT and RIGHT reactions, like the NEAR and FAR, show minima of about 830 and 890 ms, respectively, near to the head-on orientation. The elevated minimum for the LEFT response is probably due to the different response expectancies for this response: for all three observers it was the only response made with the left hand. For both the LEFT and RIGHT responses, reaction time increased rapidly and roughly linearly with rotation away from the head-on position, reaching ragged peaks of about 1150 ms at one half revolution from the head-on orientation.

DISCUSSION

We suggest the following explanations for the shapes of the three types of function, beginning with that of the TOP response. With rotation about the vertical axis, neither the position within the retinal image nor the retinal shape of the TOP stimulus key changed as the orientation of the stimulus cube was altered. Furthermore, the spatial loci in the retinal image of

all the other stimulus keys, in all orientations used in the experiment, was well separated from that of the TOP key. Therefore, observers were able to use a highly efficient and compatible spatial mapping strategy for selecting the TOP response that was independent of cube orientation, and this resulted in the fast flat reaction time function. A second strategy was also available for the TOP response: the TOP stimulus key always appeared on the cube face opposite to that bearing the V. However, this relational strategy could be expected to yield longer reaction times than the spatial strategy, and was probably not used.

As can be seen from Figure 2, the 24 stimulus cube orientations fall into four groups (corresponding to the four rows of the figure), with the same six images appearing in each group. The only difference among the four images in a single column of the figure lies in the different orientation of the V on the cube's bottom face. Thus, for the four keys on the vertical sides of the cube, the <u>only</u> cue that can be used to determine which face of the cube bears the stimulus key is the position of the V relative to the stimulus key. With respect to the V, the four stimulus keys fall into two classes, with the LEFT and RIGHT keys in one and the NEAR and FAR keys in the other.

The V exhibits lateral symmetry, so although a stimulus key appearing to the side of the V can be quickly identified as

either the LEFT or the RIGHT key, the choice between these two is not easy. The orientation of the key relative to the V must be determined in detail before the choice can be made. Subjective reports suggest that the observers imagined themselves looking from the apex of the V towards its points, and then decided whether the displayed key was on his/her left or right. The former process implies a sort of mental rotation, although it is the observer's body image that is rotated rather than the stimulus image, as was the case in the mental rotation studies described in the introduction. The possibility that observers perform some sort of mental rotation gains some credibility from the strong similarity between the shape of the reaction time functions for the LEFT and RIGHT responses and those obtained by Cooper and Shepard (1973) and by Hintzman et al (1981). However, the fact that Hintzman found evidence of mental rotation in only one of the two types of tasks he used should be borne in mind. The slight minimum at one-half revolution from the head-on position for the RIGHT response was much more pronounced in the data of one observer, who reported that in these orientations he recognized the cube as being reversed, and chose his response on the basis of the spatial strategy, and then reversed it.

With regard to the NEAR and FAR stimulus keys, the same rotational strategy was possible. That is, the observer may mentally rotate his/her body image into alignment with the V, and then make a NEAR or FAR response according to the near or far location of the stimulus key. This rotational strategy would

account for the sloping skirts of the NEAR and FAR response functions for orientations near the head-on position. However, the flat plateaus observed in the NEAR and FAR functions over much of the range of orientations probably reflects the use of a second, non-rotational strategy. The V is asymmetric about its horizontal bisector, so the NEAR and FAR keys are uniquely coded by their relation to the V: the apex of the V always points towards the NEAR key, and the open end of the V always points towards the FAR key. The coding of the NEAR key is slightly more direct than that of the FAR key, because the apex of the V almost touches the NEAR face of the cube, whereas the points of the V do not reach far enough to touch the FAR face. Secondly, the V is easily interpreted symbolically as an arrow pointing towards the NEAR key, which has a high compatibility. This means that the observer can always select the NEAR and FAR responses by using a relational strategy: if the V is pointing <u>towards</u> the stimulus key, press the NEAR button; and if it is pointing <u>away</u>, press the FAR button. Use of this relational strategy would yield reaction times that are relatively independent of cube orientation and, except for orientations near the head-on position, apparently faster than those obtained with the rotational strategy. It is not clear whether the observer first processes cube orientation in order to choose between the rotational and the relational strategies, or pursues both strategies in parallel. In the latter case, the response would be determined by whichever strategy first produced a decision,

and the observed composite reaction time functions would reflect the fact that each strategy wins over only part of the cube's rotational cycle.

In either case, the utility of relational strategies obviously depends on the cue used to mark the orientation of the cube. In the present experiment, we chose a symbol with left-right symmetry but with no top-bottom symmetry, and placed it left-right symmetrically, but top-bottom asymmetrically on the bottom face of the cube (that is, the V was placed between the center of the bottom face and its NEAR edge). A relational strategy could have been used to select the TOP response, since the TOP key always appeared on the face opposite the V. This strategy would have depended only on the <u>position</u> of the V, and not on its shape. However, this strategy was probably not used because the spatial strategy consistently gave faster decisions. A relational strategy could also be used for the NEAR and FAR responses. Moreover, the strategy could make use of both the position of the V and its shape: the V appeared on the NEAR half of the bottom face, and its apex also pointed towards the NEAR face. No relational strategy was possible for the LEFT or RIGHT responses, because the LEFT and RIGHT stimulus keys were placed symmetrically with respect to both the position and the configuration of the V. Had we chosen a different letter such as an E or an F, or placed the cue differently, then a relational strategy could have been used for these responses also. On the other hand, choice of an E would have prevented a relational

strategy for the NEAR and FAR responses based on the shape of the letter, although one based on the position of the cue would still be available.

EXPERIMENT 2

The first experiment involved rotations of the cube about its Y-axis, and we found three different types of reaction time functions for the five responses. We suggested that they could be explained as different combinations of the results of applying the spatial, relational, and rotational strategies. We now report several tests of the consistency of these hypotheses, in which we examine reaction time functions for rotations of the cube about the other two principal axes. In Experiment 2, we rotated the cube about its Z (depth) axis, and in Experiment 3 about its X (lateral) axis.

For rotations of the cube image about its Z-axis, the simplest predictions follow for the NEAR and FAR responses. Since both NEAR and FAR stimulus keys appeared at a fixed location in the image, independent of the cube's orientation, we expect observers to use a spatial strategy for selecting these responses. The location of the TOP stimulus key varied with cube orientation, but was always opposite the V, so that a simple relational strategy would select the correct response. Selection of the LEFT and RIGHT responses presents difficulties similar to those experienced in Experiment 1, when the cube was rotated about the Y-axis. The locations of the LEFT and RIGHT stimulus keys vary with cube orientation, and the lateral symmetry of the V complicates the choice between them. We expect the observer to make the choice by mentally rotating his/her body position, this

time about the Z-axis, until the face marked with the V is on the bottom.

We were also interested in determining whether rotations about the Z-axis took longer than corresponding rotations about the Y-axis. People have much more experience in relating to objects in orientations that involve rotations about the vertical axis than about other axes, owing to the ubiquitous influence of the gravitational vertical. One might therefore suspect that rotations about other axes would require more cognitive effort, and therefore take longer, than those about the vertical axis.

The method for Experiment 2 was exactly the same as Experiment 1, except that the 24 stimulus cube orientations represented a complete rotation about the depth, or Z-axis, of the cube, rather than the vertical Y-axis of Experiment 1. Note that the baseline orientation of the cube for these rotations was the canonical orientation described above, with the cube image's vertical axis corresponding to the gravitational vertical, and the lateral axis rotated 10 degrees (0.18 rad) away from the head-on position. Photographs of the 24 orientations are shown in Figure 5, with the same two caveats as before: all five stimulus keys appear in each image in the photograph, whereas in an experimental trial an observer never saw more than one key. Second, the 2-D photograph is often difficult to interpret as a 3-D image, whereas observers virtually never encountered depth ambiguity when the images were displayed with SpaceGraph.

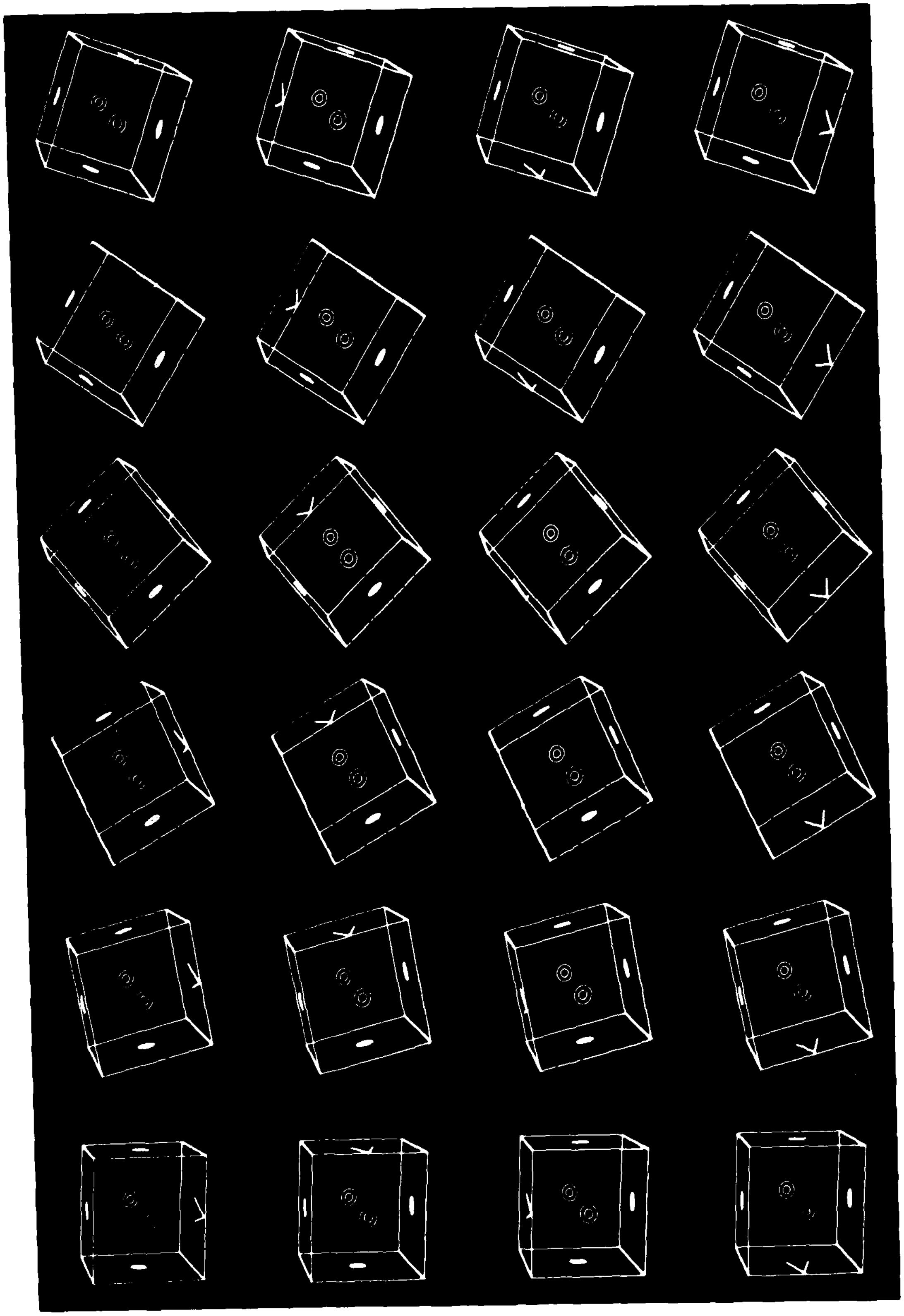

Figure 5: The 24 orientations of the stimulus cube used in the Z-axis task. On each trial, the observer saw only one of the five stimulus keys.

The same three observers served, and data were collected for six sessions, of which the first was discarded as before to reduce familiarization effects. This resulted in 60 observations per data point, 20 from each observer.

RESULTS

Errors

The three observers made 14, 20, and 9 errors respectively, for a total of 43 errors in the total of 7200 trials, an average rate of 0.6%. LEFT/RIGHT confusions accounted for 32 of the errors, and the remaining 11 were irregularly distributed among the other possibilities.

Movement times

Pooled movement times for each of the five responses are shown in Figure 6, and it can be seen that the functions are all approximately flat. The leakage present in Experiment 1 (see Figure 3) has almost disappeared, and the movement times for the TOP and FAR responses are again about 80 ms longer than the others. However, all responses are about 50 ms faster than in Experiment 1, presumably as a result of practice.

Reaction times

Pooled reaction times are plotted for each response in Figure 7. The functions for the NEAR and FAR responses are

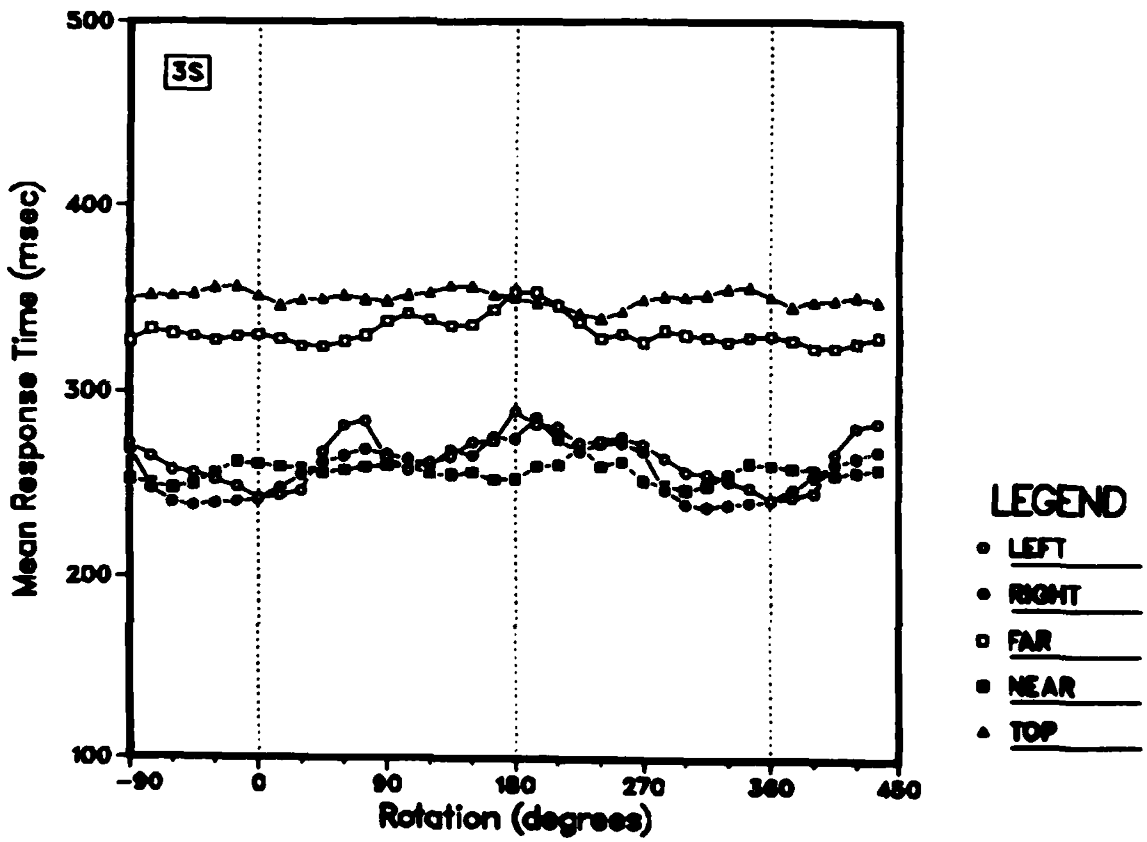

Figure 6: Mean movement times in the Z-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

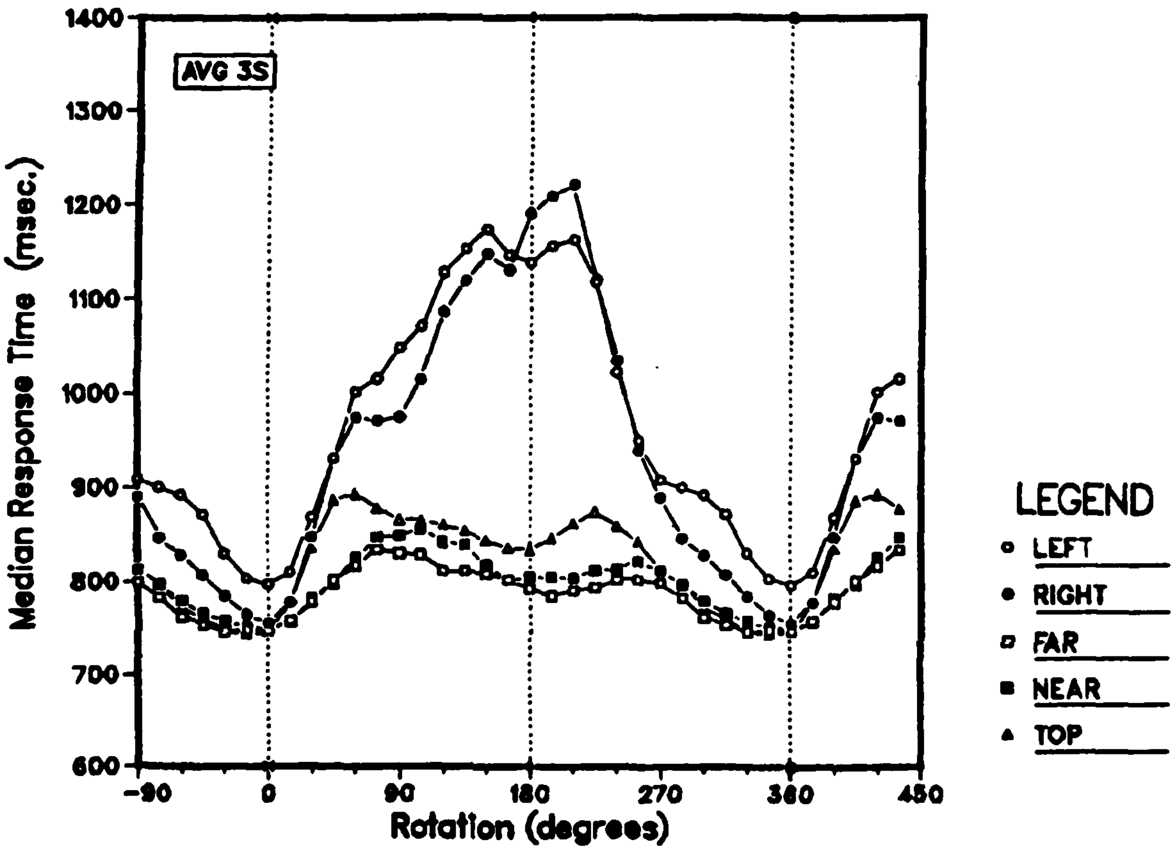

Figure 7: Mean decision times in the Z-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

plateau-shaped, and are very similar both in overall shape and level to the corresponding functions in Experiment 1 (Figure 4).

The functions for the LEFT and RIGHT responses are sharply peaked, and again are very similar in shape to those of Experiment 1. However, the reactions are about 50 ms faster at the minimum than in Experiment 1, and about 100-150 ms slower at the maximum.

The function for the TOP response is roughly plateau-shaped, but there is a minimum at one half revolution from the head-on orientation, giving the function an M-shape. At the minimum at the head-on orientation, the reaction times are equal to those of all the other responses, except the LEFT response which again is elevated by about 60 ms, perhaps because of response expectancies. Elsewhere, the TOP response function lies above the other plateau functions, but only by 20-60 ms.

DISCUSSION

When the stimulus cube is rotated about its depth axis, the retinal images of the NEAR and FAR stimulus keys remain invariant under the rotation. However, observers were apparently not able to use a 2-D spatial coding strategy directly for the NEAR and FAR keys, because, if they had, the reaction time functions would have been flat and fast like that for the TOP response in Experiment 1. Why could a spatial strategy not be used for the NEAR and FAR responses in the present experiment? One

possibility is that the spatial strategy of Experiment 1 did not <u>require</u> the image to be interpreted as a 3-D object, but could be applied directly to the raw retinal image. Thus, however immediate, automatic, and salient the 3-D percept was, the TOP response in Experiment 1 could be made on the basis of a patch of light at a particular place in the retinal image. Two aspects of the image may have discouraged or prevented use of the spatial strategy in Experiment 2. First, the NEAR and FAR stimulus keys appear in very similar positions within the retinal image (see Figure 5), so application of the spatial strategy to the retinal image requires a finer discrimination than was necessary to identify the TOP key in Experiment 1. Secondly, the main spatial separation between the NEAR and FAR stimulus keys was in the depth dimension, and apprehending this separation therefore required that the image be interpreted in 3-D before the difference in depth between the NEAR and FAR keys became available to support response selection.

The similarity of the NEAR and FAR functions in Figure 7 to those in Figure 4 suggests that observers used the same strategy for these keys in the two experiments, and we argued above that this must be a relational strategy that relied on the V pointing towards the NEAR key and away from the FAR key. As before, we attribute the sloping skirts of both functions to the use of a rotational strategy near the head-on orientation. Thus, for small rotations of the cube, it is presumably faster to rotate one's body image mentally into frontal alignment with the V and

then respond spatially, while for larger rotations it is faster to determine the relationship of the stimulus key to the V without mental rotation.

The plateau shape of the TOP function suggests that the TOP key also was identified by means of a relational strategy: the TOP key is on the face opposite the one bearing the V. A dip would occur at one half revolution if observers selected the TOP response whenever they saw a "blob" on both the upper and lower faces of the cube, without having to decide which blob was the key and which was the V.

The LEFT and RIGHT responses present the same logical difficulties to the observer as they did in the Y-axis experiment. Again, because of the lateral symmetry of the V, the observer cannot easily use a relational strategy, and must therefore rely on the rotational strategy at all orientations. As before, we observe that reaction times increase dramatically with increasing rotation away from the head-on orientation.

EXPERIMENT 3

In Experiment 3, the cube was rotated about its lateral (X) axis. Rotated in this way, the LEFT and RIGHT stimulus keys occurred in fixed and widely separated positions in the image, regardless of cube orientation. Under these conditions, we would expect an observer to use a retinal spatial strategy, which should result in fast, flat reaction time functions. On the other hand, the positions of the TOP, NEAR, and FAR stimulus keys varied with cube orientation. Since each of these keys could be directly related to the V on the bottom of the cube, we expect observers to use a relational strategy for selecting each of these responses. The resulting reaction time functions should be plateaus, probably with sloping skirts near to the head-on orientation.

The method for Experiment 3 was exactly the same as Experiments 1 and 2, except that the 24 stimulus cube orientations represented a complete rotation about the lateral X-axis of the cube, rather than the vertical Y-axis of Experiment 1. Photographs of the 24 orientations are shown in Figure 8, again with the same two caveats as before.

The same three observers served, and data were collected for six sessions (24 blocks of data), of which the first was discarded as before, to reduce familiarization effects.

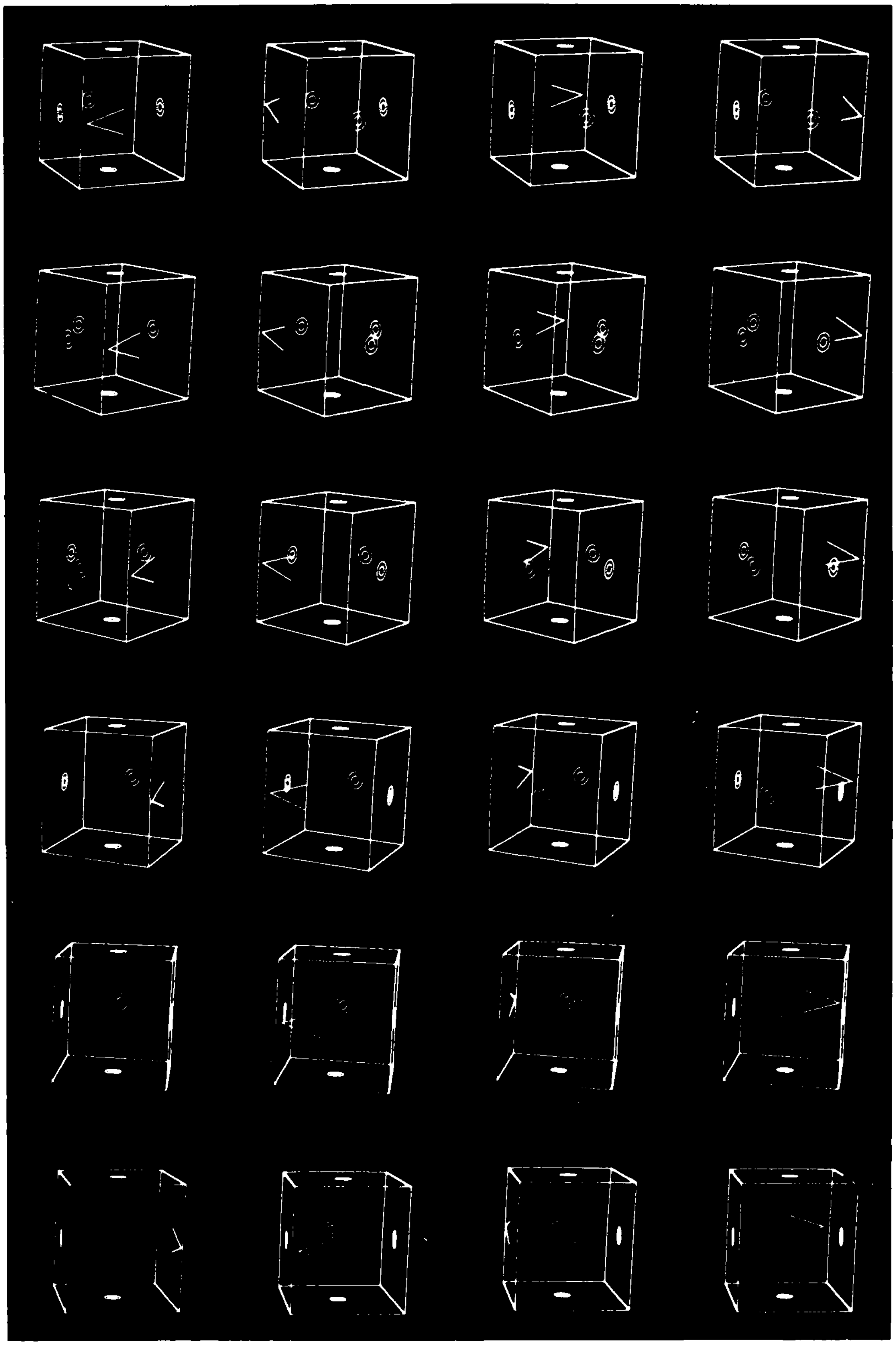

Figure 8: The 24 orientations of the stimulus cube used in the X-axis task. On each trial, the observer saw only one of the five stimulus keys.

RESULTS

Errors

The three observers made a total of 44 errors in the total of 7200 trials, for an average rate of 0.6%. One observer made 12 errors, and the other two made 16 each. Two errors were LEFT/RIGHT confusions, 22 were TOP/NEAR confusions, 15 were TOP/FAR confusions, and the remaining 5 were NEAR/FAR confusions.

Movement times

Pooled movement times for each of the five responses are plotted in Figure 9. The functions are all essentially flat, especially those for the LEFT and RIGHT responses. There appears to be slight leakage from the reaction time functions into the movement time functions for the TOP and FAR responses, with peaks at one quarter and three quarters of a revolution away from the head-on orientation.

Reaction times

The pooled reaction time functions for each of the responses are plotted in Figure 10. The functions for the LEFT and RIGHT responses are fast and almost flat, showing that rotation of the cube image about its lateral axis did not have any effect on the time taken to identify these faces. The NEAR and FAR response functions have pronounced peaks at about one half revolution from

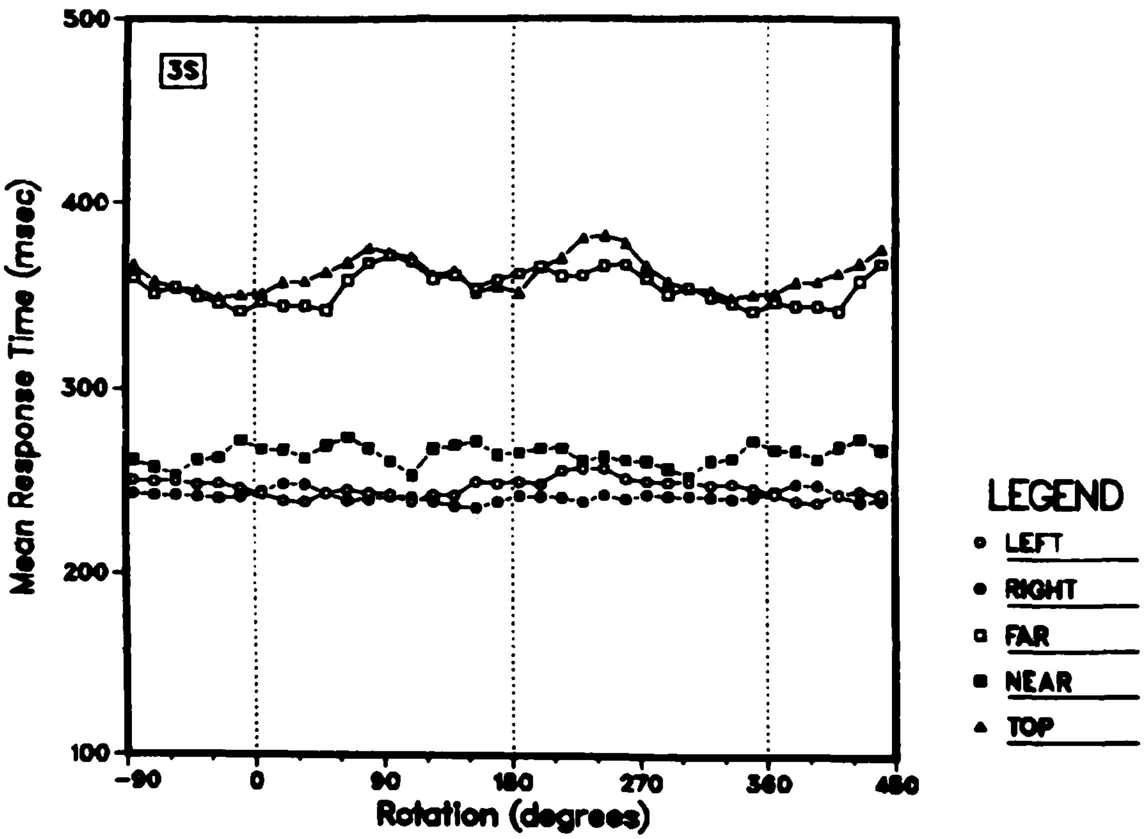

Figure 9: Mean movement times in the X-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

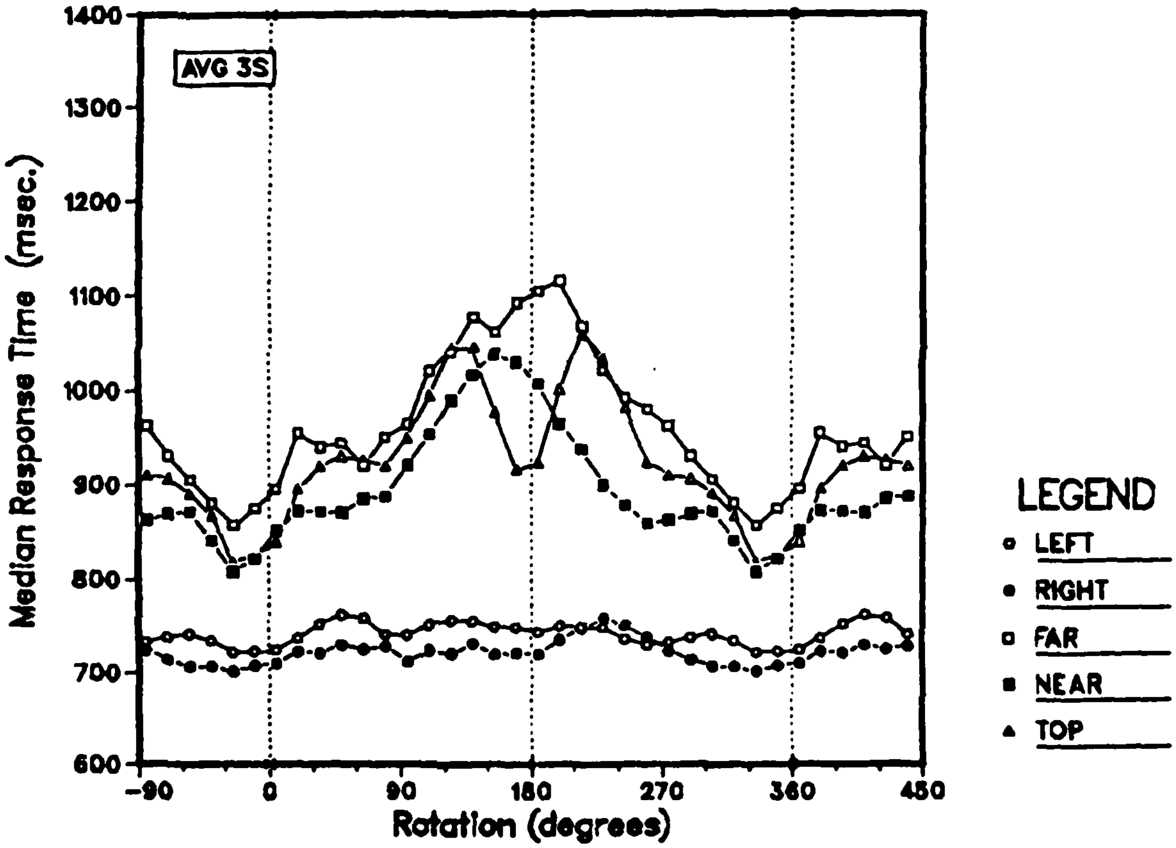

Figure 10: Mean decision times in the X-axis task for each of the five response buttons, as a function of the orientation of the stimulus cube.

the head-on orientation, where the reaction times are about 1100 ms (NEAR) and 1200 ms (FAR). For both functions, the primary minimum is slightly displaced from the canonical orientation, corresponding to the view in the bottom right image in Figure 8. The reaction times at the minima are slightly elevated relative to their values in Experiment 1 (Figure 4), and lie at about 830 ms for the NEAR response and 900 ms for the FAR response. In each function, there are subsidiary minima at about one quarter and three quarters of a revolution from the canonical orientation. The function for the TOP response is similar to those for the NEAR and FAR responses, except that, instead of a peak at one half revolution from the canonical orientation, there is a minimum. This gives the function an oscillatory character, with pronounced minima at the four quarter-revolution orientations, and pronounced maxima at intervening orientations. The minimum at the canonical orientation (870 ms) is more pronounced than the other three (970 ms) and the maxima closest to the reversed orientation (1100 ms) are higher than the others (1000 ms).

DISCUSSION

When the cube is rotated about its lateral axis, the LEFT and RIGHT stimulus keys both remain at fixed positions within the retinal image, and furthermore these positions are both near the "outside" of the retinal image, and no other stimulus key approaches these positions. Thus all the conditions required for

the retinal spatial strategy to select the LEFT and RIGHT responses correctly were met in this experiment. The flatness of the reaction time functions for these two responses, and the similarity of these times (720 ms and 750 ms respectively) to the times for the TOP response in Experiment 1 (730 ms), support the conclusion that the spatial strategy was, in fact, used. Observers' reports provide further confirmation.

The TOP reaction time function passes through four distinct maxima and minima as the cube image completes a single revolution about its lateral axis. The minima correspond to images in which the cube is seen with its edges aligned vertically and horizontally. At these orientations, the V and the TOP key appear either at the same depth but at maximally different positions in the retinal image, or at maximally different depths but at very similar positions in the retinal image. These are also the orientations where the position of the TOP key is maximally distinct from those of the NEAR and FAR keys. At the intermediate maxima, on the other hand, it is much harder to decide which face the V is on, and the position of the TOP key in the retinal image is quite close to that of either the NEAR key or the FAR key. However, the sense of this confusion is opposite to that required to explain the subsidiary maxima in the NEAR and FAR functions (see below). That is, the subsidiary maximum in the NEAR function occurs when the TOP key is close to the FAR key in the image, rather than close to the NEAR key as the explanation would require. A sinusoidal shape similar to the

shape of the TOP function is also obtained by plotting the distance between the TOP key and its nearest neighbor in the retinal image. Several of the details in the shapes of the functions may be explainable in terms of such detailed properties of the images involved, but the evidence is insufficient to warrant firm conclusions.

The functions for the NEAR and FAR responses are, with minor exceptions, very similar to those for the LEFT and RIGHT responses in Experiment 1 (Figure 4). This conflicts with our expectations for what strategies would be appropriate in the present experiment for the NEAR and FAR responses. We expected the relational strategy to be applied, since the NEAR stimulus key is always "pointed to" by the V, and the FAR key is "pointed away from." But the functions obtained are not the plateau-shaped functions we associated with use of the relational strategy. Rather, they repeat the peak-shaped functions found in Experiment 1 for the LEFT and RIGHT responses, which we ascribed to use of a rotational strategy. Comparison of Figures 2 and 8 suggests that the "pointing" aspect of the V was much easier to apprehend when the cube was rotated about its vertical axis (Fig. 2) than when it was rotated about its lateral axis (Fig. 8). In the former case, the V always lay in a true horizontal plane near the bottom of the image, whereas its position was much less predictable in the latter. This may have made it much harder to apply a relational strategy in the present experiment than in Experiment 1, leading to adoption of a rotational strategy instead.

One detail of the NEAR/FAR functions is also of interest. Both the NEAR and the FAR reaction time functions have minima (and the NEAR function also has a maximum) at orientations that are slightly displaced from the canonical orientation (the corresponding images appear at the right hand end of the bottom and the second rows of Figure 8, respectively). In these orientations the V and the three stimulus keys are most evenly spaced in the retinal image, and thus least easily confused. As in Experiment 1, the minima occur where the spatial strategy and the relational strategy both select the correct response, and the maxima occur where the relational strategy selects the correct response and the spatial strategy selects the competing, wrong response. This suggests that response times may have been affected in a minor way by spatial properties of individual images. However, we have been unable to arrive at a description of these effects that is simple enough to be testable with the present data.

EXPERIMENT 4

In typical real-life applications of a display such as SpaceGraph, objects will likely appear in arbitrary orientations. Users of the display will therefore not be able, in general, to use strategies that capitalize on properties of a particular set of orientations. In Experiments 1 to 3, for example, observers could use the fact that all orientations represented rotation about a single axis. In two of the foregoing experiments, observers were able to use a direct spatial encoding strategy for selecting one or two of the available responses, because the corresponding stimulus keys appeared at a constant position within the stimulus image for all the stimuli within the experiment. This was a consequence of rotating the cube around a single axis. In a wider context, the spatial strategy will be less effective, because it will be appropriate in only some of the orientations presented, and the observer will have to decide on its applicability before applying it.

Experiment 4 was designed to measure the speed and accuracy of choice reactions when rotation might be about any one of the three principal axes, and the observer had no prior knowledge about which would occur on a particular trial. This was accomplished by including in the stimulus ensemble orientations representing rotations about all three major axes. A subset of the 70 different orientations in Experiments 1, 2, and 3 was used in order to maintain comparability. Eight stimulus cube

orientations were selected from each of the three experiments, representing between them a complete rotation of the cube image about each of the three major axes of the cube by equal increments of one eighth of a revolution. The 24 images are shown in Figure 11, and correspond to the first and fourth columns of images that appeared in Figures 2, 5, and 8. The block of eight images at the left show rotation about the vertical, Y-axis of the cube (read row by row), and the middle and right blocks of images show rotation about the Z- and X-axes, respectively. Of the 24 images shown in Figure 11, only the 22 <u>different</u> ones were used in the experiment. The canonical orientation that appears three times in Figure 11, at the upper left of each block, was counted as a single orientation. The method was identical to that of the first three experiments, except that each experimental block contained only 110 different stimuli (5 stimulus keys x 22 orientations) instead of 120 as in the earlier experiments. The same three observers served, and six sessions of data were collected from each, with each session consisting of four blocks of experimental trials in which every stimulus occurred once. As before, data from the first session was discarded to reduce familiarization effects.

RESULTS

Errors

The three observers made 17, 29, and 21 errors respectively, for a total of 67 errors in 6600 trials, an average rate of 1.0%.

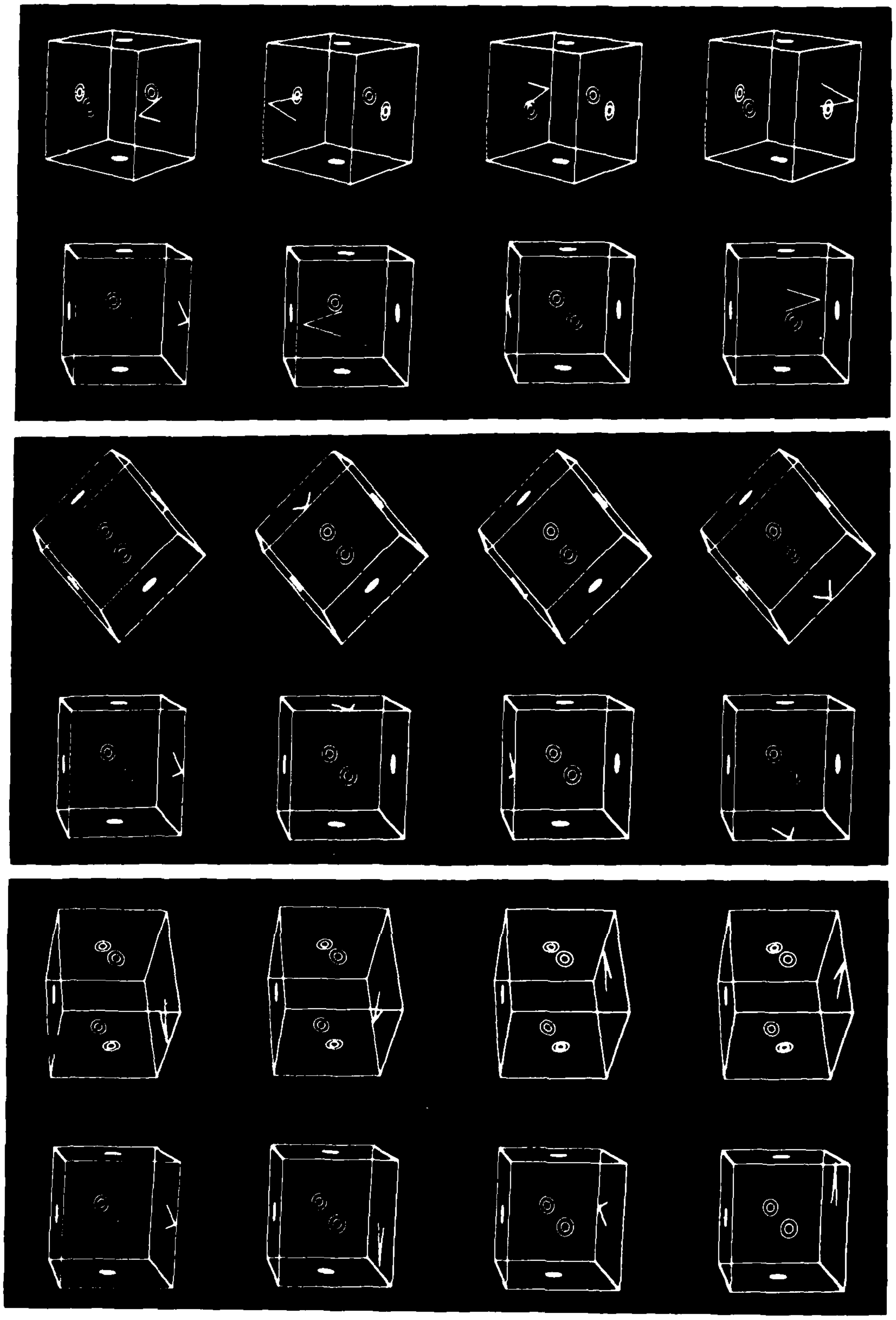

Figure 11: The 22 orientations of the stimulus cube used in the XYZ-axis task (the second and third instances of the canonical orientation, at the upper left of the middle an right blocks, were omitted). The block of eight images on the left correspond to a complete rotation about the vertical Y-axis of the cube, and the middle and right blocks correspond to rotations about the depth, Z-axis and the lateral, X-axis respectively. On each trial, the observer saw only one of the five stimulus keys.

Fifty five of the 67 errors were LEFT/RIGHT confusions, and 39 of these were made on just two of the 22 stimulus cube orientations, those that represent an inverted orientation (i.e. one half revolution away from the canonical orientation) about the X-axis or the Z-axis. The V appeared on the uppermost face of the cube image in these two orientations (and only in these two), and the spatial strategy worked for the LEFT and RIGHT responses for one of these orientations but not for the other. The only difference between the two cube images was that the apex of the V pointed towards the observer in one case, and away from the observer in the other.

Movement times

Movement times were similar to those obtained in the earlier experiments, and will not be described further.

Reaction times

The reaction time functions for the 15 combinations of the 5 responses and the three rotation axes are shown in Figure 12. The rows of panels correspond to the 5 responses, and the columns show the effects of rotations about the Y, Z, and X axes. Two functions appear in each panel. The dashed line represents the reaction times obtained in the present experiment, with "mixed" rotation axes, and the solid line represents the reaction times obtained for the identical stimuli in Experiment 1, 2, or 3, with "pure" rotation about one of the three axes. In addition, each panel is labelled with a letter to simplify reference.

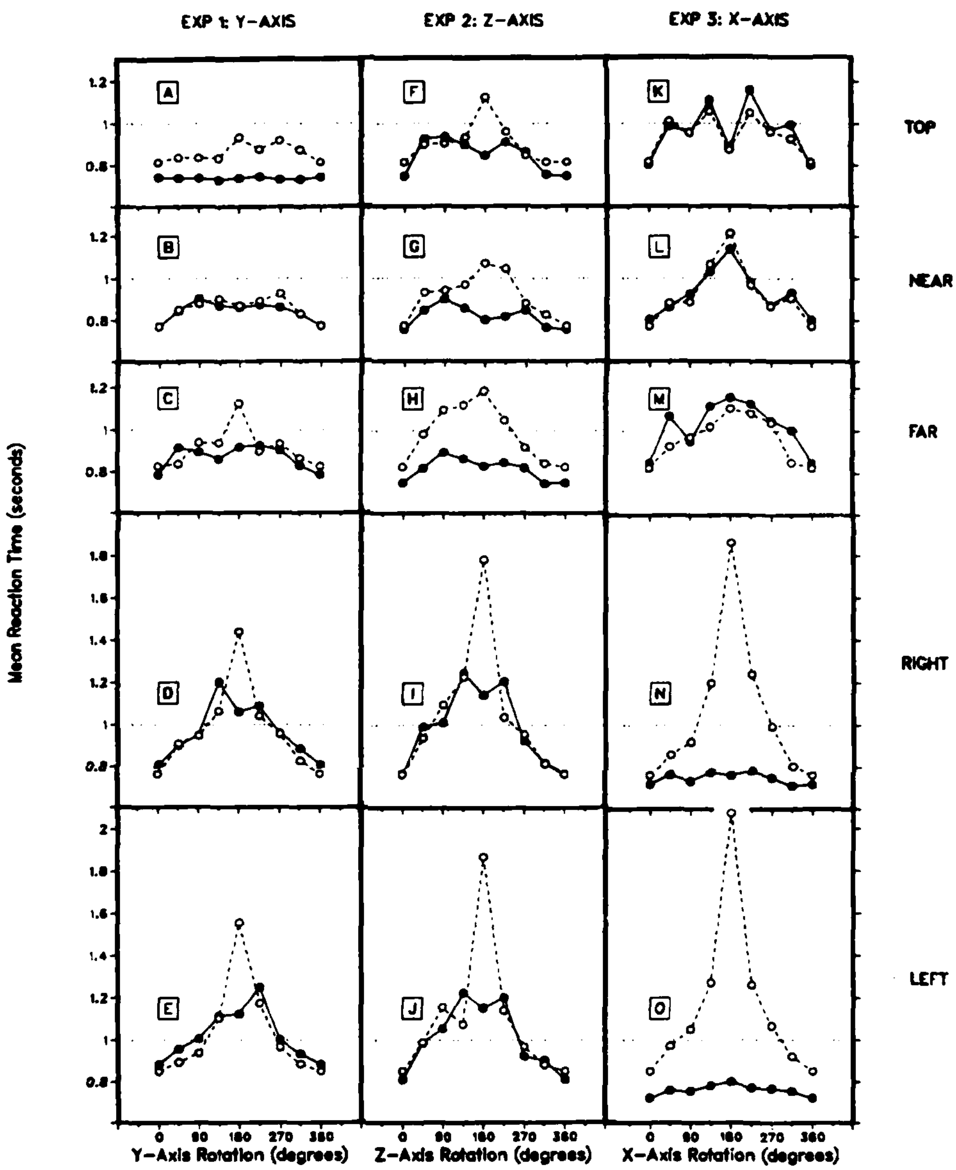

Figure 12: Mean decision times in the XYZ-task (dashed lines joining open data points) are compared with mean decision times obtained on the same stimuli in Experiment 1, 2, or 3 (solid lines joining filled data points) as a function of orientation. Functions are plotted separately for each of the five responses (rows), and for rotations about the Y, Z, and X axes (columns).

First, it is important to note that the solid functions do not correspond exactly to those plotted in Figures 4, 7, and 10, because boxcar averaging was applied to the data points that appear in Figures 4, 7, and 10, but not to those that appear in Figure 12.

There seem to be three different relationships between the results obtained with pure rotations (solid functions) and with mixed rotations (dashed functions). In four panels (B, K, L, M) there are only minor differences between the two functions. In six more panels (C, D, E, F, I, and J) there are only minor differences except at the half-revolution orientations, where the mixed functions show a marked peak. In the remaining five panels (A, G, H, N, and O), the reaction times are longer in the mixed condition (dashed function) than in the pure condition (solid function) at all orientations, the difference ranging from 50-100 ms at the canonical orientation to over a second at the reversed orientations. Of these latter five panels, three (panels A, N, and O) correspond to the conditions in which the spatial strategy could be applied in Experiments 1 and 3. The other two (G and H) correspond to conditions in which a relational strategy was apparently applied in Experiment 2. In summary, the reaction times obtained in the present experiment were quite similar to those obtained in the earlier experiments, except that (a) the peaks that occurred in the reversed orientations were much more pronounced, and (b) no fast, flat reaction time functions were obtained similar to those in earlier experiments whenever the spatial strategy could be applied.

DISCUSSION

We first focus attention on the pairs of functions for the LEFT and RIGHT responses shown in the six lower panels of Figure 12. There is a striking similarity among the skirts of 10 of the 12 functions, the exceptions being the "pure" (solid) functions in panels N and O where the spatial strategy was applied. There are two aspects to the similarity. First, setting aside for the moment the data points at the half-revolution orientations, the mixed function is similar to the pure function in panels D, E, I, and J. This suggests that, for the LEFT and RIGHT responses, observers used the same strategy in the pure and mixed conditions. Second, the mixed functions are very similar (again excluding the data point at the peak) in panels D, E, I, J, N, and O, and the pure functions are similar in panels D, E, I, and J. This suggests that the amount of rotation was more important than which axis it was about. These observations lead to a striking conclusion: that the basic strategy for determining the LEFT and RIGHT responses was similar in all four experiments, except for pure rotations about the X-axis when the spatial strategy could be applied. Because of the peaked shape of all of these functions, we believe the common strategy was rotational. The effect on reaction times of rotating the cube away from the canonical orientation was very similar for all three of the rotation axes, both for the pure and the mixed conditions. This result was quite unexpected.

The only orientations where major differences occurred between the present experiment and the earlier "pure" rotations were those one half revolution away from the canonical orientation. Here, the reaction times for the LEFT and RIGHT responses were very much longer than before. There is an obvious explanation for the peaks in the Z-axis and the X-axis functions. The two images involved were the only two in which the cube image was inverted, with the V on its uppermost face. When the cube was inverted by rotation around the Z-axis, the apex of the V pointed towards the observer and the stimulus key on the left of the image required a RIGHT response. When the cube was inverted by rotation about the X-axis, on the other hand, the V pointed away from the observer and the stimulus key on the left of the image required a LEFT response. Furthermore, this relationship between the direction of the V and the reversal of the spatial coding was the opposite of that applying when the cube was rotated about the Y-axis. That is, when the cube was inverted, the spatial strategy could be applied directly if the V pointed away from the observer. But when the cube was upright, with the V on the under face, the spatial strategy had to be reversed if the V pointed away from the observer. The fact that peaks occurred also in the mixed functions in panels D and E suggests that observers were influenced by this inconsistency. Similar factors may account for the much smaller peaks that occurred in the mixed functions in panels C and F. The observer could resolve this quandary by rotating the cube mentally about the Z-

or the X-axis, whichever was appropriate. Our observers reported great uncertainty in choosing which axis to rotate about. This indecision, possibly accompanied by unsuccessful initial rotations about the wrong axis, may well account for the considerably lengthened reaction times to the half-rotated cubes.

SUMMARY AND CONCLUSIONS

When observers are asked to identify the marked face of a static 3-D outline image of a cube, presented on SpaceGraph, and press the corresponding response key on a fixed homomorphic response cube in front of them, three strategies are available for selecting the appropriate response:

1. The Spatial strategy, which yields most accurate and rapid responses. The spatial strategy can be applied (a) for all responses when the orientation of the stimulus cube corresponds to that of the response cube, and (b) for any individual response for which the stimulus indication appears in a fixed location in the stimulus image, and this location is both distinctive and near to the edge of the 2-D retinal projection of the image. As long as these conditions are met, varying the orientation of the stimulus image has virtually no effect on the responses.

2. The Rotational strategy permits the spatial strategy to be applied indirectly, that is, preceded by a mental rotation of either the observer's body image, or of the stimulus image, whose purpose is to compensate for the differences in orientation between image and control. However, the mental rotation takes time, and the time increases with the amount of rotation required. Thus responses take increasingly longer to select with this strategy as the discrepancy between the orientation of the stimulus image and that of the fixed response cube increases.

3. The Relational strategy depends on unique relationships between the cue used to indicate the marked face and the cue provided for establishing the orientation of the stimulus image (a V drawn on the cube's bottom face in the present experiments). Thus, the viability of this strategy can be controlled by the inclusion of appropriate properties in the cues that indicate image orientation. The responses selected with this strategy are slower than those selected by the spatial strategy in the absence of rotation, but are faster than those that require substantial rotation (e.g. rotation through more than about 90 degrees).

The obtained response functions suggest that, during the course of the experiments, observers made use of all of these strategies, and learned to select for each condition the strategy that would yield the fastest correct response.

The foregoing results have some strong implications for implementations of 3-D displays. One of the major questions raised by the advent of true 3-D displays in 3-D control tasks is the perennial problem of inside-out versus outside-in view. Should the remote-controller of a moving vehicle see a display in which the vehicle is fixed and the background moves, similar to the operator's view through a car windshield or from an aircraft cockpit, or a display in which the background is fixed and the vehicle moves, such as the displays usually presented to a Ground

Controlled Approach (GCA) operator? Although direct vehicle control is usually easier from the inside-out view, there are several situations in which the outside-in view is either preferred or currently used, such as in LSO (Landing Signal Officer) and GCA tasks or when several moving objects must be simultaneously monitored or directed. Unfortunately, the outside-in view also has some disadvantages. Ambiguities of interpretation and control reversals often occur when the longitudinal axis of the vehicle is pointing towards rather than away from the operator. The results of the experiments described above suggest that observers choose to make their control decisions from the inside-out viewpoint wherever possible, because this leads to the fastest and most accurate results. This, in turn, has strong implications for what provisions should be made for the operator of a 3-D display to be able to interact with it. For example, it suggests that controls should be provided that would enable the operator to rotate the contents of the display image into a preferred orientation.

A second implication of the present results is that whenever an operator is required to make directional judgments about an object displayed as a 3-D image, the object should be made asymmetric along all its axes, so as to provide cues on which the operator can base a relational strategy. For example, if an image of an aircraft is to be displayed, the size of the tail fin should be exaggerated in order to mark both the fore-aft axis and the up-down axis of the aircraft, and in addition the left wing

should be made different and if possible more salient than the right wing, for example by showing it at greater intensity, or solid instead of open, or by adding an extra symbol such as a ring through the end of the wing. Experiments to test this second prediction are under way.

REFERENCES

Cooper, L.A., & Shepard, R.N. Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Fitts, P.M., & Seeger, C.M. S-R compatibility: Spatial characteristics of stimulus and response codes. Journal of Experimental Psychology, 1953, 46, 199-210.

Hintzman, D.L., O'Dell, C.S., & Arndt, D.R. Orientation in cognitive maps. Cognitive Psychology, 1981, 13, 149-206.

Mitchell, M.J.H. & Vince, M. The direction of movement of machine controls. Quarterly Journal of Experimental Psychology, 1951, 3, 24-35.

Morton, J. Categories of interference: Verbal mediation and conflict in card sorting. British Journal of Psychology, 1969, 60, 329-346.

Peterson, J.R. Effects of S-R coding on the rate of gain of information (Tech. Rep. No. 05823-8-T). University of Michigan, December 1965.

Sanders, A.F. Some aspects of reaction processes. In A.F. Sanders, (Ed.), Attention and performance I. Amsterdam: North-Holland, 1967.

Stroop, J.R. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643-661.

Warrick, M.J. Direction of movement in the use of control knobs to position visual indicators. In P.M. Fitts (Ed.), Psychological research on equipment design, Army Air Forces Aviation Psychology Program, Research Report No 19. Washington, DC: U.S. Government Printing Office, 1947.

Woods, C.C. Parallel processing of auditory and phonetic information in speech discrimination. Perception & Psychophysics, 1974, 15, 501-508.

OFFICE OF NAVAL RESEARCH
Code 442
TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CDR Paul R. Chatelier
Office of the Deputy
Under Secretary of Defense
OUSDRE(E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Department of the Navy

Engineering Psychology Programs
Code 442
Office of Naval Research
800 North Qincy Street
Arlington, VA 22217 (5 cys)

Communication & Computer
Technology Programs
Code 240
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Manpower, Personnel and
Training Programs
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Information Systems Program
Code 443
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Department of the Navy

Physiology & Neuro Biology
Program, Code 441
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Eastern/Central Regional
Office, ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
1030 East Green Street
Pasadena, CA 91106

Commanding Officer
ONR Western Regional Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, CA 96503

Department of the Navy

Dr. Robert G. Smith
Office of the Chief of
Naval Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Dr. Jerry C. Lamb
Combat Control Systems
Naval Underwater Systems Center
Newport, RI 02840

Naval Training Equipment Center
ATTN: Technical Liirary
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment
Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and
Evaluation Group
Naval Training Equipment
Center, Code N-00T
Orlando, FL 32813

Dr. Albert Colella
Combat Control Systems
Naval Underwater Systems Ctr.
Newport, RI 02840

Dr. Gary Poock
Operations Research Dept.
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Admin.
Naval Postgraduate School
Monterey, CA 93940

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Department of the Navy

Dr. Robert French
Naval Ocean Systems Center
Sa Diego, CA 92152

Dr. Ross L. Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 0722--Room 508
800 North Quincy Street
Arlington, VA 22217

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 340F
Washington, D.C. 20361

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D.C. 20361

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D. C. 20362

Commander
Naval Electronics Systems
Command
Human Factors Engrg. Branch
Code 4701
Washington, D.C. 20360

Department of the Navy

Mr. Warren Lewis
Human Engrg. Branch
Code 8231
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Arthur Bachrach
Behavioral Sciences Dept.
Naval Medical Research Inst.
Bethesda, MD 20014

Dr. George Moeller
Human Factors Engineering
Branch
Submarine Medical Research
Naval Submarine Base
Groton, CT 06340

Head
Aerospace Psychology Dept.
Code L5
Naval Aerospace Medical
Research Lab
Pensacola, FL 32508

Dr. James McGrath
CINCLANT FLT HQS
Code 04E1
Norfolk, VA 23511

Dr. Robert Blanchard
Navy Personnel Research and
Development Center
Command and Support Systems
San Diego, CA 92152

Mr. Jeffrey Grossman
Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Department of the Navy

CDR Robert Biersner
Naval Medical R&D Command
Code 44
Naval Medical Center
Bethesda, MD 20014

Mr. J. Williams
Department of Environmental
Sciences
U. S. Naval Academy
Annapolis, MD 21402

Human Factors Section
Systems Engineering Test
Directorate
U.S. Naval Air Test Center
Patuxent River, MD 20670

Human Factors Engrg. Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Human Factor Engrg. Branch
Naval Ship Research and
Development Center
Annapolis Division
Annapolis, MD 21402

CDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, CA 93940

Dr. Julie Hopson
Human Factors Engrg. Div.
Naval Air Development Center
Warminster, PA 18974

Department of the Air Force

U.S. Air Force Office of
Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB,OH 45433

Air University Library
Maxwell AFB, AL 36112

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks AFB, TX 78235

Other Government Agencies

Defense Technical Inform. Ctr.
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12cys)

Dr. Craig Fields, Dir.
System Sciences Office
Defense Advanced Research
Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Lloyd Hitchcock
Federal Aviation Admin.
ACT 200
Atlantic City Airport,NJ 08405

Dr. M. Montemerlo
Human Factors & Simulation
Technology, RTE-6
NASA HQS
Washington, D.C. 20546

Foreign Addresses

Director, Human Factors Wing
Defence & Civil Institute
of Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
Canada

Dr. A. D. Baddeley, Director
Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Other Organizations

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Dr. T. B. Sheridan
Dept. of Mechanical Engrg.
MIT
Cambridge, MA 02139

Dr. Arthur I. Siegel
Applied Psychol. Services, Inc
404 East Lancaster Street
Wayne, PA 19087

Other Organizations

Dr. Harry Snyder
Dept. Industrial Engrg.
Virginia Polytechnic Inst.
and State University
Blacksburg, VA 24061

Dr. W. S. Vaughan
Oceanautics, Inc.
422 6th Street
Annapolis, MD 21403

Dr. Robert T. Hennessy
NAS-National Research Council
2101 Constitution Ave., NW
Washington, DC 20418

Dr. M. F. Samet
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Williges
Human Factors Laboratory
Viginia Polytechnic Inst.
and State Univesity
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Alphonse Chapanis
Dept. of Psychology
The Johns Hopkins University
Charles and 34th Streets
Baltimore, MD 21218

Journal Supplement Abstract
Service
American Psychol. Assoc.
1200 17th Street, NW
Washington, D.C. 20036 (3cys)

Other Organizations

Dr. Christopher Wickens
University of Illinois
Dept. of Psychology
Urbana, IL 61801

Dr. Thomas P. Piantanida
SRI International
BioEngrg. Research Center
333 Ravensworth Avenue
Menlo Park, CA 94025

Dr. Babur M. Pulat
Dept. Industrial Engrg.
North Carolina A&T
State University
Greensboro, NC 27411

Dr. Richard W. Pew
Information Sciences Dept.
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Douglas Towne
Univ. S. California
Behav. Technology Lab.
3716 S. Hope Street
Los Angeles, CA 90007

Dr. Stanley N. Roscoe
New Mexico State Univ.
Box 5095
Las Cruces, NM 88003

Dr. William R. Uttal
Institute for Social Research
University of Michigan
Ann Arbor, MI 48109

Dr. Robert Fox
Vanderbilt University
Department of Psychology
Nashville, TN 37240